

COST ALLOCATIONS IN A MULTIPRODUCT NATURAL MONOPOLY:
SEPARATIONS IN TELECOMMUNICATIONS

BY

MARTIN FRANCIS GRACE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1986

Copyright 1986

By Martin Francis Grace

ACKNOWLEDGEMENT

There are numerous people who I would like to thank for helping me with my dissertation. Numerous faculty members, outside of my committee have given me a great deal of assistance such as Thomas Cooper, Edward Golding, and the members of the University of Florida's Industrial Organization seminar. In addition, my committee, Roger Blair, Virginia Wilcox-Gok, Jeffrey Harrison, Richard Romano, and especially my chairman, Sanford V. Berg, deserve a great deal of thanks. I also had the assistance of some of my peers in the graduate program. Dick Scoggins, Scott Makar, and Bob Jacob all gave me bits and peices of advice that helped me shape the final product. I would also like to thank my typist, Melissa Davis, for the time she spent assisting me and typing my dissertation. Finally, I would like to thank the Public Utility Research Center at the University of Florida for its generous financial support over the last few years.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	iii
ABSTRACT.	vi
CHAPTERS	
I INTRODUCTION.	1
II COST ALLOCATIONS IN TELECOMMUNICATIONS: SEPARATIONS	10
2.1 Early State Regulation	10
2.2 Federal Involvement	17
2.3 Federal and State Competition	22
2.4 Introduction of Competition	32
III OVERVIEW OF THE APPROACHES TO ALLOCATING COSTS	46
3.1 Introduction	46
3.2 Joint Costs and the Theory of Public Goods	47
3.3 Ramsey versus Fully Distributed Cost Prices.	58
3.4 Accounting and Game Theoretic Approaches to Allocating Costs.	65
3.5 Axiomatic Approaches to Allocating Costs	80
3.6 Conclusion.	87
3.7 End Notes	89
IV COST ALLOCATIONS WITHIN A MULTIPRODUCT FIRM: THE CASE OF FIXED PROPORTIONS.	91
4.1 Introduction: Production with One Input.	91
4.2 Fixed Proportions Model.	97
4.3 Long Run Analysis	108
4.4 Increasing Returns to Scale in One Market.	110
4.5 Conclusions.	115
4.6 End Notes.	117

V	COST ALLOCATIONS FOR A GENERAL CASE	118
5.1	Cost Allocation between Two Regulatory Jurisdictions	118
5.2	A Two Jurisdiction Model with Joint Costs	121
5.3	Long Run Results When α is Exogenous .	132
5.4	Short Run Analysis with an Exogenous Change in α	137
5.5	An Application: Welfare and Ramsey Prices	140
5.6	Welfare and Passive Regulation	144
5.7	Conclusion	152
VI	CONCLUSION AND AREAS FOR FURTHER RESEARCH .	155
	BIBLIOGRAPHY.	164
	BIOGRAPHICAL SKETCH	170

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of
the Requirements for the Degree of Doctor of Philosophy

COST ALLOCATIONS IN A MULTIPRODUCT NATURAL MONOPOLY:
SEPARATIONS IN TELECOMMUNICATIONS

By

Martin Francis Grace

May 1986

Chairman: Sanford V. Berg
Major Department: Economics

Government cost allocation procedures in telecommunications require all users of a local telephone network to pay a portion of the costs of provision of the local network. This network, characterized as a joint input into the production of local and toll calls, consists of cable and telephone poles used to connect subscribers to the central switching office.

A major premise of the dissertation is the selection of a cost allocation mechanism by an administrative agency will effect the amount of the joint input a multiproduct firm will choose to employ. Most discussions of the cost allocation problem, however, treat the cost to be allocated as a fixed cost and then allocate the cost among the markets. The approach in this dissertation is different.

In many cases the "fixed" costs are really a joint input into the production of many services. For example, the local telephone network is a fixed capital investment in the short run, but when the allocator is changed there will be a change in the choice of the size of the capital investment through changes in the derived demand for the joint input.

The government's choice of allocator is very critical. Policy-makers chose between equity and efficiency in selecting the allocator. This dissertation's purpose is to apply economic theory to the allocation problem to see how the issue might be resolved and to explore the firm's input use incentives.

Chapters II and III review the of the cost allocation literature and the history of the cost allocation schemes in the telecommunications industry. Chapter IV is examination of the incentives facing the Leontief firm when the cost allocation is changed for policy reasons. Chapter V examines the same problem for a firm with a general production function to show the effects of possible approaches to allocating costs.

Finally, Chapter VI concludes and summarizes the results. Policies altering allocations between markets can cause more than changes in the output mix as the choice of the joint input is also influenced by the new allocator.

CHAPTER I INTRODUCTION

Multiproduct natural monopolies attract attention from industrial organization economists as well as from the federal government's antitrust authorities. Concepts such as economies of scope and sustainability are used to analyze the nature of multiproduct production and to explain the incentives facing a multiproduct producer. From the standpoint of policy analysis, the government has taken an interest in constraining the power of monopolies, whether through regulation or antitrust enforcement. Recently, the government dismantled AT&T, the world's largest multiproduct firm. In addition, the government, through actions taken by the Department of Justice and the Federal Communications Commission, has set five policy goals for the telecommunications industry (see Table 1.1). The government's ability to meet these goals depends, to a great extent, on the ability to find a cost allocation methodology consistent with these goals.

The government's first goal is to promote competition in the telecommunications industry. Since the 1960s, when the FCC allowed limited competition in the interstate microwave communications market, the industry has been experiencing dramatic economic growth and technological

Table 1.1
Government's Telecommunications Goals

1. Promote competition to achieve low prices and innovation. (FCC and Antitrust Division).
2. Promote financial viability of the local telephone network. (FCC and State Governments).
3. Eliminate cross-subsidies between different services. (FCC and Antitrust Division).
4. Eliminate uneconomic bypass. (FCC)
5. Promote universal service. (Congress)

change as increased uses and users have placed demands upon the industry. The government's first goal seeks to exploit all the benefits technology and competition have to offer (Kahn, 1984).

The second governmental goal is to keep the local telecommunications network financially viable in the face of competition in the toll marketplace (Lavey and Carleton, 1983). Most analysts, including Judge Harold Green, who oversaw the dismemberment of the Bell System, believe that there are advantages to having one least-cost provider of service at the local level. While promoting competition in the interexchange market, the federal and state governments have a political and social interest in maintaining the viability of local networks.

However, with the government promoting competition in the interexchange market, its third goal of making prices reflect costs forces the elimination of transfers from toll users to local users. In the past, the transfer from toll to local (viewed by some as a "subsidy") helped to keep local rates relatively low. Through the cost allocation process, whenever a toll call was placed, the local company received a payment from the toll company to cover the local company's cost in providing the call. Through many revisions, this cost allocation factor has risen from less than 5 percent in 1947 to 25 percent in 1984. This trend indicates that toll users have been paying relatively

more for the use of the local network over time. It is recognized that the amount of the "subsidy" is very difficult to determine. In fact, some argue that in the telecommunications industry a cross-subsidy between local and toll does not exist (Pelcovitz et al., 1983).

The government, in determining the amount of the transfer in the future, will upset many customers, because removal of the subsidy could lead to increases in local rates. This development could affect the second goal of keeping the local network viable. As local rates increase, customers could choose to leave the network. There is empirical evidence to suggest, however, that this fear is a bit exaggerated since the demand for local access is relatively inelastic (Dougherty, 1984). Even so, low income customers are more likely to be on the margin and would choose to leave the network if prices increased. Regulators must address the incidence of the benefits and costs of deregulation and the AT&T divestiture.

The fourth goal of the government would allow the freedom of choice for firms (and to a much smaller degree, individuals) to leave the established telephone network if they so desire. The government desires efficient "bypass" of the network if the customer makes the decision to leave the network based on the actual costs of service and not on some arbitrarily set cost of service. Large users of telecommunications services claim that because of

regulatory cost allocation schemes set up by the state and federal governments, the larger users of toll telecommunications services pay more than their fair share of the local service costs, and thus are encouraged to bypass the network. By building their own network or sharing a network with others, these large users can save money. The problem is that if the large users decide to build their own networks, then the revenue flowing back to the local company through the cost allocation process is lost. In such situations, the local customer will bear more of the burden of paying for local service, influencing the marginal customer's choice to procure local telecommunications services.

The government's fifth and final policy goal is the promotion of universal service. The idea that everyone should have access to affordable telephone service dates to the Communications Act of 1934. Proponents of universal service argue that a subsidy to ensure universal service can be justified on the grounds of public safety, income distribution impacts, and network externalities. Realistically, it is not a policy that can be discarded even if the arguments for efficiency dictate its retirement. The policy prescription of universal service has enough adherents that it will survive, even though it may conflict with the government's other four goals.

The crux of the universal service problem is "bypass." If the alleged cross-subsidy between local and toll users is to be discontinued and the current flat rate pricing scheme with unlimited calls is maintained, some marginal users will stop subscribing to local service because they cannot afford the service. universal service would be threatened. If larger users, on the other hand, are "taxed" to support the universal service mandate, the cost of telecommunications services to the larger customer increases, causing an incentive to search for a lower cost method of communicating. Since the government's policy allows these large users to "bypass" the local network and build their own network, the burden of supporting the local network devolves on those without substitutes.

It is evident that policy-makers must choose between equity and efficiency; they face a difficult choice. The purpose of this study is to apply economic theory to the problem to see how the issue might be resolved and to explore the firm's input use incentives in light of a changing cost allocation methodology. The local telephone network, which is still a natural monopoly for the purposes of discussion, is treated as a public input into the production of toll and local services. Artle and Avernous (1973) analyzed the telephone network as a public good output and derived the Samuelsonian provision conditions for a representative individual. The analysis here

examines the local network in terms of a multiproduct firm because many telecommunications firms use the local network as an input in the production of their services. The public good quality of nonrival consumption is combined with the concept of joint costs. Efficient allocations are derived and compared to second best results for firms operating under different regulatory conditions.

The familiar Averch-Johnson (AJ) (1962) story is also examined to show the interrelationship between the cost allocation process and the differences in regulatory policies between (say) the state and federal regulators. This dissertation examines the AJ model because it is the traditional way of viewing regulatory problems, but it is the effects of the cost allocation on the input choices and output mix that are the real concern. Wellisz (1963) describes a type of regulation in the natural gas transmission industry that is based upon a cost allocation formula. Wellisz points out that regulation of profit based upon an arbitrary division of costs between different regulatory jurisdictions could have an important effect upon the choice of transmission capacity and the mix of output supplied to the various markets. This is the type of behavior this dissertation seeks to examine in the telecommunications industry rather than studying the AJ "overcapitalization" result.

The remainder of this study is divided into five parts. Chapter II is a review of the literature describing the different approaches to allocating costs. The economic theory of cost allocation, as well as the accounting theory of cost allocation, is examined. Chapter III is a historical review of the telecommunications cost allocation process, while Chapter IV examines the cost allocation process in a multiproduct firm operating with fixed proportions. Chapter V is a more general formulation of the problem for a regulated firm. In this chapter effect of two different approaches to allocating costs are examined and comparative statics are employed to show the effects of changing the allocator on the output mix and input choice of a regulated firm. In this regard, it was shown that it was possible to determine how far the allocation chosen by the government differed from the allocation that would have been chosen under Ramsey pricing. In addition, a analysis is made of "regulation by cost allocation." Regulation by cost allocation policies attempt to alter the cost allocation between markets to stem off rate increases in one market rather than the traditional rate base regulation. Using a theoretical framework described by Joskow (1974), it is show that cost allocation regulation was a viable regulatory theory for describing why the government altered the cost allocation rules during the past forty years.

Finally, Chapter VI summarizes the policy implications of the analytical framework developed in Chapter V and identifies areas for further research.

CHAPTER II

COST ALLOCATIONS IN TELECOMMUNICATIONS: SEPARATIONS

The cost allocation process in the telecommunications industry (commonly referred to as separations) that existed prior to the AT&T/Department of Justice consent decree was a relatively arbitrary costing methodology, resulting, over the last fifty years, from (1) a court mandate, (2) a desire for universal service, and (3) the conflicting interests of state and federal regulators. Thus, in order to examine the current cost allocation system, it is useful to trace the development of separations from its judicially mandated inception to its present state.

2.1 Early State Regulation

Separation of costs was not a major issue in the telephone's early years. When the telephone first became established as a viable means of communication, it was concentrated in the larger cities throughout the country. The phone was a means of talking to someone or conducting business within that city: not until the year 1881 were there any intercity, much less interstate, telephone messages.

With the development of the local exchange and the growth of the use of the telephone, interexchange connection became more and more prevalent. With little or no

interconnection there was no need for separations. However, with the development of a national network, coupled with former AT&T Chairman Theodore Vail's vision of universal service, separations became a valid regulatory concern.

The first judicial look at separations came in 1913, when the Supreme Court reviewed the Minnesota Rate Cases. The Minnesota Railroads and Warehouse Commission (MRWC) set maximum rates for freight and passenger travel for the Minnesota railroads. The railroads contended that the rates set by the MRWC upset the present freight charge arrangement, which resulted in discrimination against those cities competing for business across state lines. Second, the railroad companies felt the rate to be a hindrance to interstate commerce as well as confiscatory (Minnesota Rate Cases, 1913: 773).

In a manner that was soon to be applied to the telephone industry, the court said that state regulation was not restricted to latent federal authority, meaning that the state had a right to regulate commerce within its borders. The court also examined the issue of jurisdictional cost determination. The main thrust of the argument presented by Justice Learned Hand was that the method of cost allocation should be based not on revenues, but on use (Minnesota Rate Cases, 1913: 754-757). Fairness, in the sense that all users should pay something towards the local

network, rather than efficiency, was the focus of the decision. This usage rule has become one of the guiding criteria for telephone cost separations. To see how this "separation" of plant is conceptually possible, a simple diagram, Figure 2.1, will aid in the exposition. Suppose when subscriber A places a call to subscriber B, two types of plant are used. The telephone, or station, is used along with the local trunk to route the call to the central office where it is routed to B. The types of plant employed were the joint plant, used in the production of all calls (the subscriber station and the local connecting trunk, as well as some hardware and software in the central office). This is called nontraffic-sensitive plant (NTS). This NTS is fixed plant (i.e. does not vary in the short run). In the long run, however, NTS can increase or decrease as the demand for the use of the local network changes, so in the long run it is assumed that there is no congestion. The other plant is the equipment used solely in the production of a toll call and is traffic sensitive. For a local call from subscriber A to C, a similar route is undertaken, with the exception being that the central office (CO) routes the call through local circuits and completes the call to subscriber C. In both instances, common plant and equipment are used to complete the call. For the sake of economic efficiency, it would be best to charge a price that would cover the costs

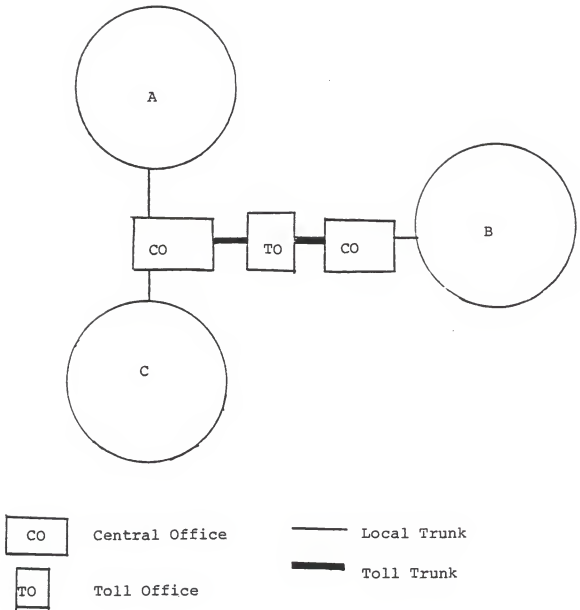


Figure 2.1
Diagram of a Call

of supplying local calls and a price to cover the provision of long distance services. With the use of joint inputs and common plant, however, this is not possible without the use of some allocation formula.

Separations, which are the cost allocation procedures in the telecommunications industry, were considered in a variety of contexts. When telephone companies were small and "city oriented," the municipalities wanted control over regulating the municipal utility. As telephone exchanges grew, the states gradually took over the regulatory function and vied for regulatory authority with the federal government.

Separations were also affected by other regulatory policies, such as those legitimized in the economic relationships within the telephone industry. By the 1930s, the regulators had studied the issues and examined the legal, economic, and institutional environment; discovered most of the important administrative problems; and set the stage for the development of a cost allocation procedure (Sichter, 1977: 11-15).

Before going further with the development of the separations procedures, a short discussion of the revenue dispersement is in order. Separations revenues generated by toll calls were returned to the local company under a scheme known as the settlements process. Figure 2.2 describes this process. Every telephone company in

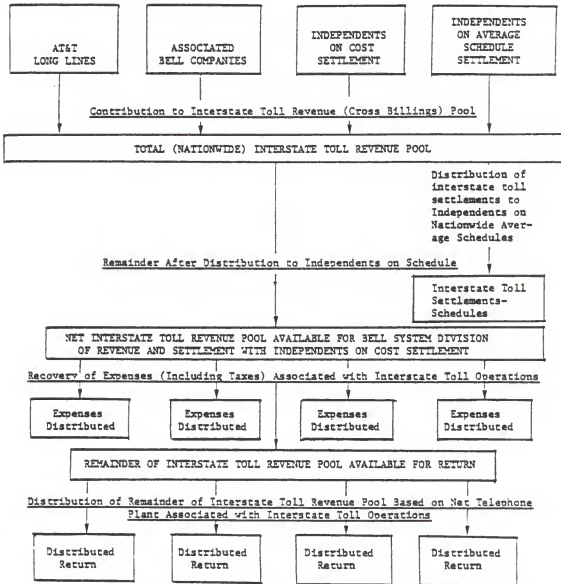


FIGURE 2.2

BREAKDOWN FOR SETTLEMENTS PROCEDURES

Source: Oettinger (1980)

the U.S. providing some portion of an interstate toll call participated in this nationwide pooling of toll revenues. Not all companies, however, were treated similarly. For small independent companies, settlements were based upon a nationwide "average cost" formula. For larger independents, such as General Telephone, there was a separate plan. This plan allowed for the independents to collect from the Bell System for the "local" costs involved in producing a toll call. The Bell Companies participated in a similar plan to cover their local expenses in producing toll calls. For the present discussion, the Bell System's division of revenues and the Independent's revenue distributing scheme will collectively be called settlements.

After billing and collecting from the long distance customers, the local company sends its toll receipts to the nationwide interstate toll revenue pool. As mentioned above, the small companies receive their payments from the pool based upon a notion of average costs for similarly situated firms. The larger independents and Bell companies were allowed to recover all their interstate expenses from the pool. After complete expense recovery, the remainder of the pool was distributed based upon the amount of the capital devoted to toll held by each company. Each company thus received a nationally averaged rate of return on plant allocated to interstate toll. Similarly, for

each state there was a similar intrastate toll settlements pool whereby intrastate revenues were distributed.

As can be imagined, this pooling agreement added some possible incentives to the firm. Expenses no longer had to be minimized because they were directly refundable from the pool. Secondly, the pooling arrangement would cause the firm to prefer to invest in toll equipment since the more toll equipment a company had, the larger would be the portion of the pool returned to the company. Although, not the focus of this dissertation, the interaction between settlements and separations has very important ramifications in the efficient input use which should be kept in mind when discussing the telecommunications industry.

2.2 Federal Involvement

The 1930s brought about a second major judicial examination of the separations issue and the regulation of interstate calls. The U.S. Supreme Court, in Smith v. Illinois Bell (1935), settled the growing controversy between two theories of separations: board to board versus station to station. In Figure 2.1, the differences between the board to board and the station to station approaches can be seen. Under board to board, a call originating with subscriber A and terminating with subscriber B would be allocated only the cost of going from A's central office (board) to B's central office (board). No local equipment

(trunk or CO) would be allocated. Under the station to station approach, all equipment used from subscriber A's phone (station) to subscriber B's phone (station) would be subject to allocation to the toll market. Along with every other company with the exception of New York Bell, the Illinois company had used the board to board method. The Bell Operating Companies (BOCs), like Illinois Bell, furnished exchange and both interstate and intrastate services. Much of their plant and equipment was used jointly for these services. The accounting practices of these companies did not provide a reasonably accurate segregation of plant investment and expenses between toll and exchange. Furthermore, the procedures did not provide for any segregation between interstate and intrastate toll services. Consequently, a comparison of the relative profitability of the BOCs' exchange, interstate toll, and intrastate toll operations could not be made directly from the accounting records used at that time (Smith, 1930: 148). Of course, the relevance of "profit rates" when capital is jointly used is questionable in any event.

Under the old license contract agreements between AT&T and the BOCs, the long distance rates were determined by the board to board procedure. The consumer paid for the use of the toll (that is, billing, operations, and collections) equipment. Rates did not include costs for the use of the exchange plant and facilities, or for

the use of plant and equipment between the customer's equipment and the exchange switchboard, for which the subscriber had paid in the monthly service charge. A distinction was made at a basic level, where toll charges were made only on the basis of plant which could be identified as being dedicated solely for toll use. To employ one of these methods over the other implies something about the firm. For example, a multiproduct firm that is partially unregulated (as was AT&T-Long Lines prior to 1939) will have incentives to put all (or a large percentage) of its capital into the regulated market. This is essentially AT&T's reaction to the regulatory incentives and is reflected in the choice of the board to board accounting practice. This point is discussed further in Chapter V.

Alternatively, the station to station method has toll rates intended to include compensation for the use of facilities and services needed in the completion of toll messages from one subscriber station to another. Under this method, exchange rates are intended to help cover costs for the use of facilities and services needed in transmitting and receiving messages wholly within the exchange area. Illinois Bell contended that the problems associated with separating plant and equipment between the two jurisdictions were too large to conquer. The court rejected Bell's reasoning and argued for the station to

station theory of separations, mandating that total costs be apportioned to each jurisdiction according to uses made of plant and equipment:

The separation of the interstate and intrastate property revenues and expenses of the company is important not simply as a theoretical allocation to two branches of the business. It is essential for the appropriate recognition of the competent governmental authority in each field of regulation [that a distinction be made among the plant used]. (Smith: 148)

The court continued by saying that it recognized the inherent problems involved in the separation of plant. Extreme nicety was not required, only "reasonable measures" need to be essential. According to the decision, the company had no right to ignore the use to which the plant is put. The station to station theory, however, examines the costs involved in making a call and allocates the costs according to equipment use. For example, when a person makes a toll call, local equipment (that is, the telephone and central office equipment) is used as well as long distance facilities. Station to station would apply some costs to the local equipment and some to the toll plant, while board to board completely ignores the use of local equipment in the provision of the toll service. The decision thus had the effect of lowering local exchange rates since plant was being transferred to the long distance market. Less exchange plant allowed the firm to decrease local rates, since there was a reduction in allocated costs. In addition, since the interstate

jurisdiction had allocated more plant and equipment to its rate base, it would have to raise its prices to cover costs (assuming inelastic demand).

The Smith decision was not actually applied to the separation procedures by the Bell System until 1943, and it took an additional seven years before AT&T changed its position on intrastate toll tariffs to comply with Smith. Gable claims that this 20-year delay since the Smith decision and its implementation was made possible by the lengthy legal and technical construction of the telephone costing operation. But Temin and Peters (in press: 21) claim that AT&T was reluctant to change sooner because of the concern that its plant would be regulated by more than one utility commission and it was afraid that possibly conflicting regulation could leave some plant out of the rate base.

With the Communications Act of 1934, the Interstate Commerce Commission, which previously had authority over interstate toll rates, but which did little or no active work in the area, passed its jurisdiction on to the Federal Communications Commission (FCC). As the FCC started to tangle with regulating interstate rates, the separations issue began to reveal its complexities in greater detail. The state public utility commissions were afraid of a growing FCC that could start to regulate intrastate rates. The state's fears increased the juris-

dictional competition, which hindered the eventual separations process in later years. Also, there were definitional problems concerning how capital, plant, and expenses were to be allocated. Given the institutional, political, economic, and administrative problems arising from the issues, the separation process evolved in an erratic manner.

2.3 Federal and State Competition

Rate disparity (that is, the difference in interstate and intrastate rates for a similar call) became a major problem for the states during the 1940s and 1950s. Most of the common plant and equipment were assigned to the state jurisdiction, which meant that the costs of a call would be greater if it was made intrastate in contrast to interstate. The interstate call would have the benefits of lower allocated costs. When the FCC saw long distance earnings on the rise as a result of lower costs and new technology, it would cut the rates on interstate toll calls instead of reallocating plant to the interstate jurisdiction (viewed by the states as a more equitable position).

In 1941, the FCC investigated Bell's interstate toll rates. The interstate-intrastate rate disparity and the clash between the states and the FCC increased as a result of the investigation and order, which resulted in a substantial reduction in interstate rates (Gabel, 1967:

37). The state commissions were taken aback by a perceived disregard for their needs by the FCC. The states' ratepayers were faced with higher exchange costs and intrastate toll calls, while the price of an interstate toll call was decreasing. The negotiated interstate rate reductions only made matters worse for the state commissioners. For a similarly placed toll call, one originating and terminating intrastate, and another of the same distance originating in state and terminating across state boundaries, the difference in cost to the caller increased. The states were under pressure to resolve this perceived disparity and, as a result, the National Association of Regulated Utilities Commissioners (NARUC) cooperated with the FCC in the investigation of separations problems.

As a result of this agreement, two investigations were opened in the early 1940s: Docket 6328 and Docket 6468. Docket 6328 again brought up the related problems of station to station and board to board methods of separation and was just a rehash of the issues. Docket 6468 was another matter entirely. The FCC saw AT&T's Long Lines profits increasing, so in 1943 the FCC and AT&T negotiated a tariff reduction estimated to cost Long Lines \$50 million. The net effect of the tariff revision was to transfer \$22 million of revenues per year to the BOCs. This policy switch by AT&T, even though the company still

supported the board to board procedures, pointed to AT&T's realization of the changing importance of intra- and interstate toll service. Either through foresight or fortune, AT&T's policies reflected the exploitation of the long distance network and the increased demand the nation's businesses would place upon long distance services.

The competition to protect consumers that existed between the state public utility commissions and the FCC created more and more problems. The FCC ignored the state proposals on separations and did not take an active stance on separations, preferring the less time-consuming negotiation method for rate cases. This policy vacuum created by the FCC allowed Bell, in effect, to determine its own principles for separating and assigning plant and equipment to the two jurisdictions. The situation came to the point where changes were not made unless AT&T granted its approval (Gabel, 1967: 46, and Temin and Peters, 1985a: 20-25).

Post-War Developments

After the Second World War, the FCC and the states were able to effectively obtain uniform (on a per-mile basis) interstate rates through a \$100 million interstate rate reduction, although the BOCs had applied for increases in local rates of \$242 million. The competition between the FCC and the states again affected the division between

interstate-intrastate toll revenues. The problem was that toll earnings were repeatedly increasing at such a rate that Long Lines was forced to assent to reductions while the additions to intrastate plant required rate hikes in the local jurisdiction. This type of arrangement could have conceivably continued without further investigations into cost allocation or relative use. Because of the pressures facing the states, NARUC persuaded the FCC to try to alleviate the problem of separations according to the rules set in Smith (Gabel, 1967: 51).

In 1947, the NARUC-FCC Joint Board Separations Manual was completed. The manual, although a good first step, was riddled with problems. Because of conflicting jurisdictional interests, the manual contained procedures that give consideration to a myriad of operating conditions as well as regulatory requirements" (Porter, 1979: 7-11). The FCC never officially adopted the manual, and since NARUC is not a legislative or administrative body, little actually came from the manual's first edition. Gabel said that it is "nearly impossible to proceed with actual separations and allocation of plant or telephone property using the manual as a sole guide (1967: 57). In recognition of the weakness of the document, many plans were proposed during the next twenty years.

The 1947 plan had as its basic tenants the division of plant between exchange (local) and interexchange toll

(both intra- and interstate) use, and assignment of the plant to the proper jurisdiction where possible. If, however, plant could not be directly assigned, a conceptual construct called Conversation Minute Miles (CCM) was used as a way of assigning costs. The CCM, which was a ratio of time and distance per unit of circuitry, was to be used as a measure of relative use. The manual also identified two types of jointly used plant: The first were facilities used interchangeably and intermittently to complete intra- and interstate calls. This plant was assigned to the interstate jurisdiction through the CCM, so intrastate rates could be brought down. The second type of jointly used plant consisted of facilities such as telephone poles, cable, and conduits which "support or contain circuits, some of which were used in furnishing a single service" (private line or long distance service) and were assigned directly to the interstate jurisdiction (Gabel, 1967: 64). So, the cost allocation methodology seemed more to reflect state pressures than principles of cost causation.

By 1950, due to technological change and economies of scale, Long Lines' earnings were again growing compared to intrastate earnings. Under the threat of rate reduction by the FCC, Bell proposed a major revision of the method of allocating interexchange plant. Under Bell's proposal, called the Phoenix Plan, the book cost of toll line plant

would be allocated to the respective jurisdictions by a message-minute-mile (MMM) of the BOC circuitry to the MMM of Long Lines circuitry within the state. The ratio of the interstate MMM to total state and interstate MMM would be applied to the book cost of toll line plant. Given this description, Bell's plan can be expressed symbolically. Let MMM_{ij} equal message-minute-miles in jurisdiction i by company j , where

i = N (National average),
 S (State average),
 I (Interstate average),
 j = BOC (Bell Operating Company),
 LL (Long Lines division of AT&T),
 B = Book Cost for BOC plant used in jurisdiction,
 IR = Interstate requirement for MMM, and
 c = Fraction of cost going to interstate

Then $IR = MMM_{I,BOC} + MMM_{I,LL}$,

where $MMM_{I,LL} = \frac{(\text{Circuit miles of toll plant in BOC}) \times}{MMM_{N,LL}}$;

and $cB = MMM_{I,BOC}/IR$,

where cB = Amount of cost assigned to the interstate jurisdiction.

Given this plan, and the FCC's reluctance to make any alterations, nothing was done except to pursue another investigation into telephone toll rates. In 1951, NARUC petitioned the FCC on behalf of its state constituents calling attention, once again, to the fact that the Bell System exchange and state toll rate had increased more than \$400 million annually, while interstate toll rates were reduced by FCC action in 1943, 1944, 1945, and 1946. NARUC complained the perceived inequity was resulting from

political jurisdictional competition, and asked once again for modification of separations procedures. Simultaneously, the Senate Majority Leader, acting in his capacity as chairman of the Senate Commerce Subcommittee on Communications, sent a letter to the FCC pressing for action "at once" (Gabel, 1967: 68-69, and Temin and Peters, 1985a: 23).

The FCC again examined the separations problem and, as a result of negotiations, instituted the Charleston Plan. The affect of the plan was to shift an estimated \$90 million of plant and investment and \$22 million of related expenses to the interstate jurisdiction. The Charleston Plan had three major rules. The first was to identify plant units and assign them to specific categories. These categories (reduced from 16 to 3 for simplicity) were then to be apportioned to the services (that is, local exchange or toll) on various bases. In contrast to the 1947 plan, in place of direct assignment of local interoffice trunks to exchange service, the plant and equipment were put into a pool and apportioned on the basis of relative use. Toll connecting trunks, previously lumped with toll investment, were also placed into the same category as exchange trunks and apportioned jurisdictionally according to relative minutes of use (Gabel, 1967: 76). The end result was to shift fixed costs to the intrastate jurisdiction on a usage sensitive basis.

As a result of the plan, the interstate rate revisions reduced the interstate-intrastate toll rate disparity (the rate difference between a toll call paced interstate and one placed intrastate) by \$68 million, from \$160 million to \$92 million. The changes (which were actually increases in short haul interstate rates by \$22 million) helped to lessen the disparity. The FCC had hoped that the states would benefit from the revisions and would in turn be able to lower the intrastate rates or keep them stable. Not one state reduced its toll or exchange rates (Gabel, 1967: 77-78). According to Sichter, there were numerous reasons why the state commissions did not act upon the intrastate rate reductions. Primarily, the states doubted the ability of the toll services to bear the additional costs imposed by the station to station theory. Second, even though toll rates seemed to be based on toll "costs," as defined by the Uniform System of Accounts, to have gone beyond the accounting system would have brought out the practical and administrative problems of detailed and arbitrary cost allocations that the BOCs were trying to avoid. Finally, the total effects of toll separations were "bounded by neither the Bell System nor the jurisdictional limits of the state," creating a much more difficult problem than did the regulation of "relatively self contained exchange systems" (Sichter, 1977: 21).

The Charleston Plan only slowed the pace of intrastate rate increases. Even with the transfer of investments to interstate operations of plant and expenses, rate applications in the intrastate jurisdiction continued. By May 1953, the interstate-intrastate toll disparity returned to where it stood before the changes brought about by the Charleston Plan (Gabel, 1967: 81). It is important to realize, however, that there was no economic reason why interstate and intrastate toll rates had to be uniform. Routes that were heavily used could have different prices than routes with relatively low usage. Telephone rates for intrastate tolls could also be higher, simply because the population density is so small for areas such as Wyoming or Nevada, as compared to areas such as Connecticut or New York. The higher the population density the lower would be the per capita cost of service. During this time frame, however, regulators at the state level were not concerned with these issues. The states were concerned more with the lack of uniform pricing than efficiency.

NARUC, the FCC, and AT&T got together once more to propose, debate, table, and propose again alternative methods for separations. Accepted was a Modified Phoenix Plan (MPP), an alteration of Bell's previously proposed plan (rejected by the FCC). The MPP provided that book costs of long lines plant terminating in each state be combined with the BOC toll plant in that specific state.

The combined total investment would be apportioned on the basis of relative use (message-minute-miles). AT&T estimated the MPP would transfer \$162 million of investment and \$18 million of expenses from state to interstate operations, reducing state toll costs by 22 percent (Gabel, 1967: 82).

Symbolically, it is possible to relate the concepts in a straightforward manner:

LL = Long Lines plant (book cost) terminating in-state.
 BOC = BOC toll plant (book cost) within state
 T = Book cost for total toll plant in-state
 MMM_i = Message-minute-miles in jurisdiction i,
 where $i = I$ (Interstate)
 $i = S$ (Intrastate)

C_i = Cost assigned to jurisdiction i.

$T = BOC + LL$

$C_i = MMM_i \times T \quad i = I, S,$

where C_I are the costs that can be assigned to the interstate jurisdiction using the MMM_I for interstate calls, and C_S are the costs which can be assigned to the intrastate jurisdiction using the MMM_S for intrastate toll calls.

The Modified Phoenix Plan seemed to stop the interstate toll reductions by the FCC, since more capital and expenses were assigned to the interstate jurisdiction, reducing pressures to lower interstate rates and increasing intrastate rates. The MPP lowered intrastate revenue requirements by \$35 to \$40 million, which amounted to 6.5

percent of the combined Bell System intrastate net operating income in 1956 (Gabel, 1967: 91)

2.4 Introduction of Competition

In the early 1960s, the FCC decided in the Above 890 (1962) decision to open up microwave channels to the use of industry and government, permitting microwave long distance service in competition with Bell's own Long Lines division. Bell, reacting in an uncharacteristically swift manner, introduced Telpak and WATS services to preserve its place in the market. Gabel notes Bell's increasing awareness of separations and its affect upon the Company's competitive well-being: Bell resisted changes that would adversely effect its position in the intercity toll market.

As a result of changing technology and perhaps a change in demand for long distance services, Long Lines' earnings once more began to increase. With pressure from NARUC and the FCC, AT&T started to introduce "After 9 Rates" in the interstate market on an experimental basis. These innovative rates allowed the customer to call after 9 p.m. and receive cheaper station to station rates. The price for a station to station call was reduced in order to make better use of transmission facilities at off-peak times. The various BOCs participating in the experiment anticipated decreasing revenues, but after implementation, revenues actually increased: regulators and managers were introduced to the concept of demand elasticity (Gabel,

1967: 100-103). By 1964, NARUC again was proposing amendments to the MPP Separations Plan. NARUC was pressuring for a program to reallocate local dial central office equipment (LDCOE), extended area service (EAS) equipment, and to reclassify telephone service into timed and untimed minutes of use. AT&T was cool towards the amendments, so nothing was accomplished until 1965, when NARUC and the FCC proposed what was to be called the Denver plan.

The Denver plan provided that EAS and inter-office trunks were to be assigned directly to the exchange category. Toll connecting trunks were to be apportioned on the basis of conversation minutes of use and the book costs of subscriber lines were to be apportioned by a user use factor (UUF), which can be formulated as

$$UUF = SLU + (R_1 * R_2)$$

where

$$\begin{aligned} SLU &= \text{Subscriber line usage factor} \\ &= \frac{\text{Interstate Minutes of Use}}{\text{Total Minutes of Use}} \end{aligned}$$

$$\text{and } R_1 = (\text{Toll Users}) / (\text{All Users});$$

$$R_2 = \frac{\text{Interstate Toll Messages}}{\text{Total toll messages}}$$

AT&T claimed the Denver agreement increased interstate revenue requirements by \$98.5 million, thereby reducing intrastate accordingly. At the time, the nationwide toll disparity had grown to \$200 million. Gabel asserts that there should have been a decrease in intrastate rates by

the states, but after 14 months, only 17 states had reduced rates to allow their ratepayers to benefit from the plan. Gabel again points to the fact that revenues from the reduction in revenue requirements went to AT&T-Long Lines (Gabel, 1967: 121). Figure 2.3 shows some interesting evidence as to why intrastate rates did not fall. In Figure 2.3, the operating ratio for the combined BOCs is shown with the corresponding Long Lines ratio. The operating ratio is defined as the ratio of operating expenses to operating revenues. In a loose sense, it can be used to indicate a profit margin. During the period 1965-1969, the BOCs' operating ratio remained fairly constant while Long Lines' ratio actually fell, seemingly as a result of the increased revenue requirements. Intrastate rates did not fall. This reluctance to lower rates could be due to the increased expenses even though some capital costs were shifted to Long Lines. Previous to 1965, the operating revenue was increasing and it is possible that the increased expenses during this time period kept the BOCs from reducing rates.

In 1969, the matter of separations came before the FCC once again. This time, however, a major reform occurred which resulted in the 1971 Ozark plan. Prior to the plan, the FCC allowed AT&T to lower long distance prices; the states were once again faced with the prospects of intra-

state rate increases. This time the states went to Congress for help (MacAvoy and Robinson, 1983). Legislation was proposed to remove the FCC's jurisdiction over separations. The FCC reacted by proposing a Joint-Board, composed of representatives of NARUC and the FCC. Congress then modified the legislation to give jurisdiction over separations to the Joint-Board and enacted the modified jurisdiction into law. The Ozark plan, the result of the FCC-NARUC Joint-Board, was another attempt to allocate costs in a politically acceptable manner, going into greater detail and utilizing more ratios and usage factors.

The plan involved a number of major innovations. Interstate costs, according to the Ozark plan, are a function of the interstate SLU, subscriber plant costs, and a ratio of nationwide costs to local costs for a toll call. Toll costs were allocated on the basis of a new construct, the subscriber plant factor (SPF):

$$SPF = 0.85SLU + 2SLU*CSR$$

where

SLU = Subscriber line usage factor, and

CSR = Local average for a three minute station to station call as a percentage of the nationwide industry average for a three minute station to station call at the nationwide average length of haul.

The additive SPF with its numerous parts was designed to recognize the fact that price affects quantity demanded.

Toll calls are priced according to distance and time, so the use of the composite station ratio (CSR) has the effect of increasing the assignment of costs to the interstate jurisdiction as the length of time and the distance of the haul increases. Obviously, the problem of efficiency remains with the new formula, because the Ozark plan still allocates costs based on use, rather than on cost causation. Efficiency plays only a small role in the process.

The SLU factor was modified to reflect the deterrent effects of distance of haul and length of time. Since price for toll calls is now a function of distance and time, the deterrence inherent in toll calls becomes an implied function of price. So as distance and length increase, toll prices rise, which is an efficient way of pricing except that, as mentioned above, the method of determining the deterrent effect is based upon usage and not cost causation. For local exchange equipment, as before the Ozark plan, all equipment is allocated on the basis of relative use. Toll minutes of use, however, are weighted by an arbitrary factor of 2 to reflect the fact that the cost per toll minute of use is greater, on average, than the cost per exchange minute of use, while the 0.85 term reflects the cost of subscriber plant per toll minute of use.

The plan provides for the separation of all equipment into two categories: non-traffic sensitive (common plant)

and traffic sensitive plant. Non-traffic sensitive plant (NTS) would be allocated by the use of the SPF, so $SPF \times NTS$ is the interstate cost assignment of joint capital. The traffic sensitive portion (approximately 75 percent of plant) would be allocated on a relative minutes of use basis, as under the Denver plan, but a variable weighting factor is used to reflect variable costs per minute of the toll call, depending on the types of equipment used in the office. Older equipment such as electromechanical switches would be given a higher weight since that type of equipment costs more to operate than the newer digital switching equipment. This could have the result of diminishing the incentive to purchase new equipment, since the claim on the settlement pool diminishes. In addition, the settlements process gave additional revenue to firms with relatively high operating costs. On the other hand, higher settlements went to the firm with relatively larger investment. So the role of settlements complicates the incentives facing the firm.

The SPF that was the result of the 1971 Ozark plan was the final version of two FCC decisions in 1967 (Docket 16258) and 1969 (Docket 17975). This evolution reflects the increasing importance of political tradeoffs in contrast to economic efficiency considerations. The separations procedures were designed to allow a transfer

from the interstate toll market to the local exchange market.

Table 2.1 shows the allocative changes that have occurred over time for various classes of plant and equipment. As can be seen, the amount allocated to the interstate jurisdiction has risen in every category except the last two. The last two categories contain plant that is allocated to the toll network with some being allocated to interstate and the residual to intrastate. As an aside, it is interesting to note that this allocation to interexchange plant has remained constant since 1957.

Table 2.2 shows the effects resulting from the change in the separations formula on interstate revenue requirements and the weighting factor for each plan. As the plans have become more complicated than just relative usage (SLU), the allocations have multiplied the effects of the SLU. Figure 2.4 shows how the SLU and the newer plans have increased the allocation to the interstate jurisdiction over time.

Equipment allocated to the two different jurisdictions was classified by the regulators as nontraffic sensitive plant (NTS). In theory, this NTS corresponds to the fixed costs of the local network. The NTS includes items such as telephone poles, cable from the central office to subscribers, and certain hardware at the central

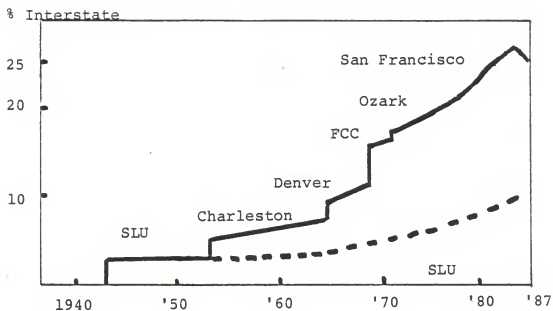


Figure 2.3
Interstate Allocation over Time

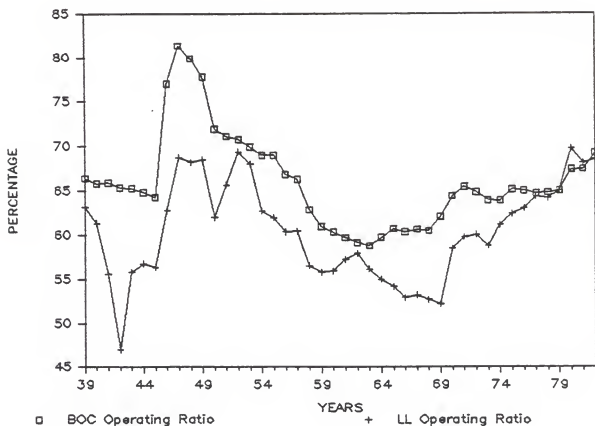


Figure 2.4
Operating Ratios Over Time

Table 2.1
Allocation Formulae/Changes

DATE	Formula/Change
1947	$SLU = \frac{\text{Interstate Toll Minutes}}{\text{Total Minutes of Use}}$
1952	SLU, (more plant assigned to NTS)
1956	Message Minute Miles
1962	Reclassification of Plant
1965	$SLU + \frac{\text{Toll Users}}{\text{All users}} \times \frac{\text{Toll Messages}}{\text{All messages}}$
1971	$SPF = SLU(2CSR + .85)$

Table 2.2
Effects of Changing Separations Formula on
Interstate Revenue Requirements and Weighting Factors

DATE Weight	Separations	Change	Increase in Revenue Requirement
1947	Original Plan		\$13 million SLU
1952	Charleston		30 million 1.8*(SLU)
1956	Mod. Phoenix		40 million *
1965	Denver		134 million 2.5*(SLU)
1969	FCC Plan		108 million 3.2*(SLU)
1971	Ozark		131 million 3.3*(SLU)

*no data available

Source: Sparling (1980: 92) and Leggette (in press: 11).

office used to produce both local and toll calls. The NTS, also subject to some arbitrary assignment, was then allocated to the local and long distance markets based on the relative use allocating mechanism.

This procedure would probably have become more complicated and inefficient if it were not for a decision by the FCC in 1969 to allow entry in Microwave Communications, Inc. (MCI), in the private line market between St. Louis and Chicago. The FCC granted MCI's petition (over AT&T's strenuous objections), but met an FCC denial when MCI petitioned to offer a service to the general public which would compete with AT&T. It was not until 1977 when the Court of Appeals for the District of Columbia Circuit reversed the FCC's decision in what has been called the Execunet decisions.

Competition had arrived in the industry, but with competition came more cost allocation issues. For example, how much should the other common carriers (OCCs) be charged for the use of the local network and interconnection. The OCCs argued that since they received an inferior grade connection compared to the connection given to AT&T, they should get a lower charge than AT&T for interconnection. These charges were negotiated among the various parties and the FCC and the resulting agreement was authorized by the FCC as the Exchange Network Facilities for Interstate Access (ENFIA) agreement.

Even with the ENFIA agreement, MacAvoy and Robinson (1983) claim the FCC was slow to implement modifications to the separations procedures to reflect the reality of competitive entry. The OCC's were charged a portion of AT&T's joint costs under the ENFIA agreement, but the charges were still less than comparable with AT&T's contributions. In addition, the FCC had mandated that costs be allocated on a fully distributed costing scheme, which has a tendency to cause prices to be higher than marginal costs (and perhaps higher than the average costs of potential entrants).

In 1980, the FCC undertook another investigation of the separations formulation in Docket 80-282. In addition, the Department of Justice brought suit against AT&T for antitrust violations. Before the FCC could finalize its results of the separations investigation, the Department of Justice and AT&T settled the antitrust suit. As a result of the settlement, however, the cost allocation problem still did not go away, and because of the concern over universal service and local rates, cost allocation took an even more important role in the structurally altered industry.

Figure 2.4 shows that the effect of the allocation mechanism has changed dramatically over time in transferring costs from intrastate to interstate jurisdictions, especially in the last 15 years. As a result of Docket

80-282, the FCC-NARUC Joint-Board, in early 1982 imposed the San Francisco plan on the industry. Unlike previous plans, no new allocation procedure was developed. Under this new plan the interstate NTS allocation would be frozen (or capped) at its 1981 level with a gradual movement towards a nationwide figure of 25 percent over the next eight years. According to the Bell System Separations Statistics for 1983, the SPF allocation differs widely from state to state, ranging from 13 percent in the area of Kentucky served by Cincinnati Bell to 62 percent for Nevada. The average for the Bell System as a whole for this time period was approximately 26.5 percent. States such as Nevada will lose settlements from the interstate toll revenue pool previously devoted to keeping local rates lower while states such as Kentucky will gain additional funds under the new 25 percent rule.

The FCC has recently proposed access charges to replace the separations process. The states have not been enthusiastic about the access charge arrangement proposed by the FCC. NARUC, in fact, went to court to stop the implementation, while other groups went to Congress to get the arrangement legislatively altered. Congress, although interested, has not acted, and the courts so far have allowed the FCC to keep much of its plan intact. The net result is that whether or not the states, telecommunications companies, and other interested parties like it,

the industry is going to replace separations with access charges.

Before adopting a new cost allocation scheme in the guise of access charges, there is a need to understand how allocation schemes affect capital formation and overall industry efficiency. The remainder of this dissertation will examine further aspects of cost allocation. Before examining the effects of allocation on the firm, a theoretical overview of the proposed cost allocation mechanisms is presented in Chapter III.

CHAPTER III OVERVIEW OF THE APPROACHES TO ALLOCATE COSTS

3.1 Introduction

Before examining the literature on cost allocation it is necessary to ask an important question: Why allocate costs? An economist's response would be that costs do not need to be allocated. Firms do not care how costs are allocated. A profit maximizing firm, whether competitive or monopolistic, equates its marginal revenue to its marginal costs. Since the market determines prices, cost allocation interferes with the maximization process, introducing inefficiencies. As a positive science, economics is concerned with efficiency, and arbitrarily allocating costs can lower efficiency.

Others, such as state and federal legislators, utility regulators, and consumers, temper their desire for efficiency with their desire for fairness. The allocation of costs is fraught with a lack of a workable definition of fairness. All utility commissions are charged with setting "fair" rates, but each state or federal agency can define a different fairness that can, in turn, have different distributional and efficiency results. The U.S. Congress, in enacting the Communications Act of 1934, decided that universal service was a desirable result, but little did it realize the ramifications in determining the method of

implementing Universal Service. Courts and administrative agencies at both the federal and state level are wrestling with this conflict between fairness and efficiency.

This haphazard regulation of costs resulting from the Communications Act and subsequent interpretation is now conflicting with the forceful presence of competition. Customer classes, which in the past were satisfied with the implicit taxes and subsidies within the telecommunications industry, are employing alternative methods of communications that avoid the traditional network and do not pay any "taxes" to support universal service. Large users are bypassing the telecommunications network because it is less costly for the firm to do so. This is an implicit realization that the cost allocation process was no longer "fair" to them.

The reason that cost allocation mechanisms are politically important is because society prefers equity balanced with efficiency. Cost allocation schemes must provide a level of both equity and efficiency that will satisfy a large group of users. A priori this is a difficult task. This chapter will survey approaches to allocating costs to show how the schemes attempt to satisfy either or both of these goals.

3.2 Joint Costs and the Theory of Public Goods

The cost allocation problem was first analyzed in the public utility framework at the turn of the century, when

railroad pricing was a pressing issue. The concept of joint costs was developed to explain why differential prices among classes of service were not necessarily discriminatory and were in fact needed to ensure the survival of the railroads.

The local telephone network, like the railroads, is used to supply many services (such as local calls and toll services) to many customer types. The optimal allocation of the costs among the various customer classes can proceed under either of two routes: joint costs or as a public good input. Either route, however, as will be shown below, yields the same result. The concept of joint production has a very long history in economics, dating to the seminal article by Taussig (1891) and brought into focus in the Taussig-Pigou debate in the Quarterly Journal of Economics, (1913). They debated the relationship of costs to services in the railroad industry and whether service for different customers or uses (transportation of coal versus transportation of copper) was homogenous. Taussig asserted that price differentials were attributable to joint costs. Pigou, on the other hand, claimed that there were common costs of railway services (where relative usage determined cost causation); Pigou reasoned that since railroads had the ability to price discriminate between buyers, the price differentials were discriminatory.

Their disagreement centered on their respective views of the cost characteristics of railroads (common versus joint). If the services are homogenous, then different prices for these homogenous services would be discriminatory. But if the services were distinct and required a joint input, then different prices could be charged to reflect the different demands. Each service would then be contributing different amounts to cover the joint costs. Although Pigou and Taussig did not directly address the issue of cost allocation, they did set the stage for further analysis.

In a prophetic article, Sichler (1928) described the telephone network as being characterized by joint costs. Sichler also realized the implications of joint costs in terms of the cost allocation problem:

Ultimately, the rise in cost resulting from a large area covered or a larger number of subscribers will bring about a condition where increased costs will have caught up with or passed the utility of certain [customer] classes . . . and the question of how to allocate costs and [set] rates will come forward as a serious problem. (1928: 175)

Sichler describes the joint cost problem in terms of the residential customer's and the business customer's use of the local telephone network. Each customer type places a demand upon the network for service. Correspondingly, in a more modern framework, this would also include the demand placed on the local network by toll users whether the users are business or residential toll customers, or

interexchange carrier customers. Each customer classification places a distinct demand for a distinct service upon the local network for interconnection and can be charged a distinct price to cover the joint costs.

To illustrate this result, one can use Marshall's (1927) stylized mutton and wool joint production model. Sheep are used to produce two products: mutton and wool. Each of these products has a demand curve which, when appropriately indexed and vertically summed, yields the aggregate demand curve for sheep. The intersection of the aggregate demand curve with the supply curve for sheep determines the equilibrium amount of sheep used to produce mutton and wool, with the price for mutton and wool being set by the marketplace--reflecting the demands for the respective product. Figure 3.1 shows this case graphically. Let D_m and D_w reflect the demands for mutton and wool, respectively. The vertical summation of the curves yields the aggregate demand curve D . The intersection of the aggregate demand curve and the supply curve for sheep (S_s) gives the equilibrium quantity (Q_e) of sheep. Prices p_m and p_w are thus determined by the relative importance of the respective demands towards the equilibrium output Q_e . The total of the prices $p_m + p_w$ thus covers the costs of producing sheep for mutton and wool. A method of cost allocation based upon these distinct demands thus assures a firm's viability, all other things being equal.

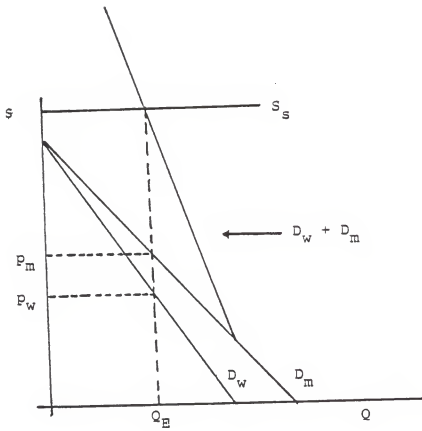


Figure 3.1
Marshall's Example

The analysis of joint production presented above is very similar to that for the efficient provision of a public good. The idea that the telephone network experiences joint production is well accepted (Kahn 1971: 172), but the concept of the local telephone network as a public good input has not been fully developed. For example, if the local telephone network could be defined as the wire and telephone poles from business and residents to the central office, and the portion of the central office which is essentially a fixed cost (i.e., it does not vary with output), this local network can be used as an input to produce many services. If a customer desires information from the telephone company, the customer can use the local network to get directory assistance. If the customer desires connection with another local subscriber, the customer can use the local network for interconnection. The local company thus uses the local network to provide these local services, each of which has a value to customers. The local network is also used as an input in the production of toll calls by interexchange carriers.

In order to connect a telephone in Gainesville, Florida, to one in Manchester, New Hampshire, the toll company must use two local networks to produce the call. Each of these local networks can be a public good input for local as well as toll services. Since the local network has many users, one more user will impose almost zero costs

on a network already in place. Ideally, this implies there are no congestion costs, but if there were congestion, additional costs could be imposed on the cost causers, or a larger network with no congestion could be built. Because of the ability to reduce congestion by building, and charging for, a larger network, it is assumed that there is no congestion. If the local network is treated as a public input into the production of various services, there should exist a derived demand curve for the local network for each of these services. The vertical sum of these derived demand curves equated to the marginal cost of building a "unit" of the local network will give the efficient size for the network. This is the same story presented in Figure 3.1.

Ellickson (1978) provides a simple analytical example of the equivalence between a joint production equilibrium and the equilibrium attained in the efficient production of a public good, or Lindahl (1959) equilibrium. A Lindahl price reflects the price that an individual consumer would pay for a unit of a public good. Unlike competitive private good provision which requires a uniform market-determined price and various quantities demanded which depend upon the individual's demand, a Lindahl price is individualized while the quantity provided is uniform (the same for everyone). Thus, competitive private good provision requires the price to be constant while quantity

demanded varies and the efficient provision of public goods requires prices to vary while quantity demanded is held constant.

To show the equivalence of the two concepts, let tastes be described by the following utility function:

$$u_i = q_i K_i \quad i=1,2. \quad (3.2.1)$$

where q_i is the quantity demanded of a private good (q_i) and K_i is the quantity demanded of the public good for the i th individual. In addition, a feasible output allocation is defined as (q_1, q_2, K_1, K_2) so that $(K_1, K_2, X) \in Y$, and where $X = \sum_{i=1}^2 (q_i - q_i^e)$, q_i^e is the initial endowment of the private good, and Y is the technology set. The variable X thus represents the net input of private goods into public good production. The technology is linear (fixed proportions) and is described by the following expression:

$$K + q_1 + q_2 \leq q_1^e + q_2^e. \quad (3.2.2)$$

In the Lindahl equilibrium, each individual faces a price, p_q , for the private good q , but faces an individualized price, p_K^i , for the private good. By letting p_q be the numeraire (that is, $p_q = 1$) and maximizing u_i subject to the budget constraint $q_i^e - p_K^i K_i - q_i = 0$, the "pseudo" demand functions are derived for the public good. Ellickson calls these demand functions "pseudo" demand functions

because the demand curves are unobservable, unlike demand curves for private goods. To get the efficient level of public goods provision, the Samuelsonian condition (that $\Sigma \text{MRS}_i = \text{MC}$) is used to get the Pareto efficient level of the public good. The above utility maximization yields a demand for the public good

$$K_i = q_i^e / 2p_K^i \quad i = 1, 2. \quad (3.2.3)$$

Solving for price gives

$$p_K^i = (q_i^e / 2K_i).$$

The vertical summation of these demands gives the expression

$$1 = \Sigma p_K^i = \Sigma (q_i^e / 2K) \quad (3.2.4)$$

where $K_1 = K_2$ because both consumers must consume the same amount of the public good. The Lindahl equilibrium is determined where this aggregate demand curve intersects the supply curve. For the sake of simplicity, assume the supply curve is horizontal and equal to one which is the marginal cost of providing K . The Pareto efficient level of production of the public good is

$$K^* = (\Sigma q_i^e / 2) \quad (3.2.5)$$

and each consumer of the public good pays a Lindahl price

$$p_K^{i*} = q_i^e / \sum q_i^e \quad (3.2.6)$$

So, the Lindahl equilibrium yields a Pareto efficient output and a set of personalized prices.

For the joint cost example, suppose once again that there are two individuals, one of which desires mutton and no wool while the other desires wool and no mutton. Let

K_1 = Total amount of meat produced by the economy and consumed by individual 1.

K_2 = Total amount of wool produced by the economy and consumed by individual 2.

q_i = Amount of sheep kept by individual i .

With the same utility functions from equation (3.2.1), the same feasible allocation and technology set, the competitive equilibrium can be derived by maximizing utility subject to the budget constraint in equation (3.2.2). This yields the same demand functions presented in equation (3.2.3). These are the observable demand functions for mutton and wool, respectively. The aggregate demand for sheep is determined by the vertical summation of the two demand curves (equation 3.2.2). The equilibrium level of sheep production (and slaughter) is thus determined by the intersection of the aggregate demand curve and the marginal cost of production and slaughter. This process yields a price for wool and a price for mutton (equation 3.2.6), which together must cover the total costs of producing sheep.

As Samuelson (1969) has pointed out, there is no general equivalence between a joint production equilibrium and a Lindahl equilibrium due to the free-rider problem. If there are no free-riders or when each user of the public good reports its marginal valuation truthfully to the provider of the public good, then the two results are identical. For example, prior to the AT&T divestiture, the local network was owned and managed by a company involved in producing the network for other telecommunications services. Accordingly, the managers of the unified company could determine the marginal productivity for the use of the local network and construct a network based upon that knowledge. Absent perverse regulatory incentives, the problems of strategic behavior or false reporting would not arise since the managers of the network are the providers of the network.

As a note to this discussion, the existence of the free-rider is not the problem facing the managers of a firm, but it is the principal-agent relationship that would cause many of the inefficiencies in provision and use of the public input. It is very costly to get the local managers to carry out AT&T's corporate will (Shavell, 1977). This principal-agent relationship, rather than the free-rider problem, introduces the inefficiencies when the providers are the users of a public input, even with unitary ownership.

A cost allocation methodology that involves either joint cost prices (following Marshall's approach) or prices based upon a division's marginal evaluation for the local network (following Lindahl's approach) is the first best approach yielding a Pareto efficient result. Actual implementation of these pricing rules is another matter entirely. Because of the complexity or impossibility of implementation, other methods, such as Ramsey pricing rules, have been advocated to price services which have large amounts of fixed or joint costs.

3.3 Ramsey versus Fully Distributed Cost Prices

Ramsey (1927) indirectly proposed the first formal procedure for allocating fixed costs across product lines or customers. His formulation resulted in the burden of overhead costs being placed on the markets with relatively inelastic demands. Ramsey originally proposed this idea as a way of taxing goods in such a way that distortions caused by the imposition of the tax were minimized. Ramsey optima are second-best alternatives; but even so, they are often used as a benchmark for evaluating performance. The basic premise of the Ramsey result has alternatively been stated as (1) maximizing the sum of producer and consumer surplus subject to the provision that the firm breaks even; (2) determining the set of prices from which it is not possible to change in a way that permits gainers to compensate the losers (assuming

nonlinear pricing); and (3) maximizing the level of satisfaction of any one individual given the utility level of every other individual (Baumol and Bradford, 1970). In the case of a regulated multiproduct firm, an alternative characterization reflects the essence of the cost allocation problem when there are non-directly allocatable fixed costs: Minimizing the joint cost of producing each output subject to the constraint that prices derived ensure that revenues equal costs, that the firms meet demand at the given price, and that output is positive in all markets where the firm produces. For example, suppose there are two outputs and that the cost function can be written as

$$C(q_1, q_2) = K + c_1 q_1 + c_2 q_2,$$

where K is the fixed cost (overhead) or, alternatively, the public input, and c_i reflects a (positive) constant marginal cost of producing q_i . Next, assume the demand for q_i is independent of the demand for any other q_j , and can be written as

$$q_i = q_i(p_i).$$

If the firm were required to produce at a price equal to marginal cost in both markets, revenues would fall short by K dollars of the point where the firm made zero economic profits. The problem, then, is to allocate K over both markets so that the firm's budget is balanced. As Sharkey (1982b) has pointed out, there are several ways to derive

the Ramsey price, one of which is to produce outputs in the same proportion as outputs would be if prices were set at marginal costs at that output level. Therefore, p_i are Ramsey prices if

$$\frac{q_i(p_i)}{q_i(c_i)} = \frac{q_j(p_j)}{q_j(c_j)}.$$

From this result, it can be seen that if the demand for good 1 is less elastic than the demand for good 2, then the optimal price, p_1 , will have to be higher than c_1 by an amount proportionally greater than that for the second good. Accordingly, the p_i 's are Ramsey optimal if

$$\frac{p_i - c_i}{p_i} = \frac{-v}{e_i} \quad \text{for } i = 1, 2,$$

where e_i is the elasticity of demand for good i and $v > 0$ is chosen to insure a balanced budget. If $v = 0$, prices are equal to marginal costs, and if $v = 1$, revenue minus costs are maximized. The reason that Ramsey pricing is important in the area of cost allocation is that even though the Ramsey outcome is itself only second best, it is a useful benchmark for comparing other schemes.

For example, Braeutigam (1980) examined three fully distributed costing (FDC) methods which were being used by regulators at both the federal and state level to set prices in regulated industries. Prior to that, no systematic analysis existed to show the inherent inefficiencies of the rules. Braeutigam examines the case where

the FDC rules are determined for a firm that is just breaking even. Braeutigam assumes that the firm acts to minimize the cost of producing the breakeven output and that there exists an independent inverse demand function for each service such that revenue for the i th service can be written as $R^i(q_i)$. Also, the revenue contribution above directly attributable costs for the i th service is $Q^i(q_i)$, where

$$Q^i(q_i) = R^i(q_i) - C^i(q_i) \text{ for all } i = 1, \dots, n.$$

$C^i(q_i)$ is the cost of producing input i and n is the number of services or outputs. The three rules Braeutigam considers are among the class of rules that allocate the common costs, K , among the n services. Each service, i , has a fraction, α_i , of K allocated to it. Since the cost of K is supposed to be "fully distributed" among the n services, it is necessary that

$$\sum_1^n \alpha_i = 1. \quad (3.3.1)$$

Each service must generate revenues ($R^i(q_i)$) that are sufficiently large to cover both the directly attributable costs and the arbitrarily allocated portion of the overhead costs, K . Thus,

$$R^i(q_i) \geq \alpha_i K + C^i(q_i); \text{ for all } i. \quad (3.3.2)$$

The first rule requires that the α_i are proportional to gross revenues

$$\alpha_i^R = R^i(q_i) / \sum R^i(q_i) \text{ for } i = 1, \dots, n.$$

If the allocations are based on directly attributable cost, then the pricing rule is

$$\alpha_i^C = C^i(q_i) / \Sigma C^i(q_i) \quad \text{for } i = 1, \dots, n.$$

Finally, if common costs are distributed according to the relative level of outputs, then

$$\alpha_i^Q = q_i / \Sigma q_i \quad \text{for } i = 1, \dots, n.$$

Now, assume the FDC rule is being determined for a firm that is just breaking even. Braeutigam examines this case because a comparison can be made to the Ramsey outputs. Rearranging equations (3.3.1) and (3.3.2) gives the revenue contribution above directly attributable costs, or

$$Q^i(q_i) \geq \alpha_i K; \quad \text{for all } i. \quad (3.3.3)$$

When π equals zero, then

$$\Sigma Q^i(q_i) = \Sigma \alpha_i K = K. \quad (3.3.4)$$

Equations (3.3.3) and (3.3.4) imply

$$Q^i(q_i) = \alpha_i K. \quad (3.3.5)$$

Therefore, when $\pi = 0$, the FDC rule must satisfy

$$\frac{Q^i(q_i)}{Q^j(q_j)} = \frac{\alpha_i}{\alpha_j} = \frac{q_i}{q_j} \quad \text{for all } i, j. \quad (3.3.6)$$

Thus, an allocation by relative output levels requires

$$1 = \frac{Q^i/q_i}{Q^j/q_j} = \frac{p_i - C^i/q_i}{p_j - C_j/q_j}$$

requires that the ratio of the difference between price and average costs attributable to each service be identical.

Similarly, by examining the other two rules (which happen

to be identical), a zero profit rule requires the ratio of a price to average attributable cost to be equivalent for all services, so that

$$p_i/(C^i/q_i) = p_j/(C^j/q_j), \text{ for all } i, j.$$

In order to examine the inefficiency of the FDC rules, Braeutigam relates the attributable costs used by the FDC rules to the marginal costs required to generate efficient prices. By letting σ_i be the elasticity of scale for product i , then:

$(p_i - C^i/q_i)/(p_j - C^j/q_j) = (p_i - \sigma_i c_i)/(p_j - \sigma_j c_j) = 1$,
for the relative output rule, where c_i is the marginal cost of i . For the attributable costs and gross revenue rules:

$$\frac{p_i}{p_j} = \sigma_i c_i / \sigma_j c_j \text{ for all } i, j, \quad (3.3.8)$$

where $\sigma_i = c_i/C^i$ and is the elasticity of scale for good i . A Ramsey optimum requires that:

$$\frac{(p_1 - c_1)e_1}{p_1} = \frac{(p_2 - c_2)e_2}{p_2},$$

where e_i is the elasticity of demand for i . It is easy to see that FDC pricing rules will yield results different from the preferred Ramsey results, since FDC rules are based on attributable costs and not marginal costs. Therefore, compared to the second-best outcome, FDC prices increase misallocations.

A pricing scheme that would be preferable to both FDC and Ramsey pricing rules is one which would take into

account the firm's marginal evaluation for the use of the common input. This method will be discussed further in section 3.4, but to give some closure to this section, a brief discussion is in order. Assume that the zero profit constraint is an average cost pricing rule requiring the firm to price at average cost in both markets. Second, assume that the cost function has constant returns to scale. Then, by dividing equation (3.3.3) by equation (3.3.4), the following expression is derived:

$$\alpha_i = \frac{Q^i(q_i)}{\sum Q_i(q_i)},$$

where $Q^i(q_i) = R^i(q_i) - C^i(q_i)$. The variable $Q^i(q_i)$ represents the quasi-rents to division 1 which can be used to cover the cost of the common input. So, α_i depends upon the percentage of rents made by division i as a ratio to the total amount of rents. With the assumption of constant returns to scale, this α represents the firm's marginal evaluation of K (given that $C^i = c_i$), the sum of which should equal the marginal cost of providing the common input. As will be shown in the next section, the α selected in this manner is the welfare maximizing cost allocation mechanism and thus is superior to both FDC and Ramsey rules.

3.4 Accounting and Game Theoretic Approaches to Allocating Costs

3.4.1 Optimal Allocations: Lindahl pricing within a Firm

Accountants are faced with the responsibility of actually implementing the regulator's wishes to apportion joint costs among the firm's output. This has resulted in a demand by the accounting profession for ways to allocate costs. A number of articles have appeared in the accounting literature to fulfill that need. It is interesting to note that economists have been among the participants in the discussion. For example, Cohen and Loeb (CL) (1981) were among the first to bring the public good literature into the cost allocation paradigm. CL assume that a firm has a decentralized organizational structure with n divisions. Let K be a vector of input demands provided by the corporate headquarters where $K \equiv (K_1, K_2, \dots, K_n)$. The profits of the i th division less the costs of K_i provided by the headquarters will be defined as $\pi_i \equiv \pi_i(K_i)$. The term $C(K)$ is the cost to the headquarters of providing K . Since K is a public good, the management will determine a level of K to supply given the division's demands for K . It is important to note that while $C(K)$ is incurred by the headquarters under a decentralized decision-making structure, division i chooses the level of input K_i under the structure. CL also assume that the cost function is non-separable, which means that it is cheaper for a number of divisions to use the public input in a coalition than it

is for each division to go out and purchase the input separately. These conditions imply that costs are sub-additive. More formally, this can be represented as

$$C(\Sigma K_i) \leq \Sigma C(K_i) \quad \text{for } i = 1, 2, \dots, n. \quad (3.4.1)$$

The firm's overall profits can thus be given as

$$\pi(K) = \Sigma \pi_i(K_i) - C(K). \quad (3.4.2)$$

A full cost allocation can be described under CL's formulation as an n-tuple $(\alpha_1, \alpha_2, \dots, \alpha_n)$, where $0 \leq \alpha_i \leq 1$ for all i , and $\Sigma \alpha_i = 1$. For a given cost allocation, the i th division's profit less the cost of the allocated input is

$$Q_i(K) = \pi_i(K_i) - \alpha_i C(K). \quad (3.4.3)$$

So, each division's profits are a function of what has been allocated to the other divisions' profits, but total profits remain

$$\pi(K) = \Sigma Q_i(K_i). \quad (3.4.4)$$

Since the headquarters knows $C(K)$, then it is sufficient for the headquarters to be able to choose profit maximizing allocations for the input and also for it to know the marginal revenue functions for each division. The derived demand for input i is thus defined to be

$$Q_i'(K_i) \equiv \delta Q_i / \delta K_i, \text{ for all } i.$$

Since K is a public input, the management will provide only one level of K ; thus $K_i = K_j = K$. The firms' profits now take on the following form:

$$\pi(K) = \Sigma Q_i(K) - C(K).$$

Now suppose that there exists some profit maximizing amount of K called K^* , then at the maximum

$$\Sigma Q_i'(K^*) = \delta C(K^*)/\delta K. \quad (3.4.5)$$

At the optimum, the sum of the divisions' marginal revenue productivities must be set equal to the marginal cost of provision the public input. Thus, an optimal cost allocation mechanism can be defined in a manner analogous to a Lindahl equilibrium. From (3.4.5) the optimal allocation is

$$\alpha_i^* = Q_i'(K^*)/\Sigma Q_i'(K^*). \quad (3.4.6)$$

This result states that each division should be charged based upon the division's proportion of marginal profitability at the optimal level of input use. Equation (3.3.4) above stated the same result for a constant cost industry where the optimal allocation mechanism depends upon the proportion of quasi-rents generated by the division to the total amount of quasi-rents.

3.4.2 The Nucleolus Scheme

Hamlen, Hamlen, and Tschirhart (HHT) (1977) describe other methods accountants have proposed to allocate costs, one of which is called the nucleolus method (Schmeidler, 1969). The idea behind the nucleolus method is similar to the Rawlsian notion of social justice. The nucleolus criterion for a cost allocation mechanism requires that the minimum surplus ($e_{\min}$) of any one division must be maximized among

the set of N divisions in the firm. Surplus for the coalition of divisions ($S \subseteq N$) is defined as

$$e_S = \sum K_i - v(s) \text{ for all } S \subseteq N,$$

where $v(s)$ represents the value or benefit to coalition S of using the input K . Accordingly, the maximization problem takes the following form:

$$\max e_{\min} = \min \{e_S\} \text{ for } S \subseteq N$$

and the surplus for any coalition S represents an index of a coalition's objection to the payoffs its members are receiving in the grand coalition. The smaller the surplus, the greater will be the objection. A benefit to this nucleolus scheme is that its result is always in the core which implies pareto efficiency. The core of a game or economy is the set of allocations that cannot be improved upon without making anyone worse off. This will be discussed further in the section concerning game theoretic approaches to allocating cost.

3.4.3 The Moriarity Approach to Allocating Joint Costs

HHT describe another allocation method proposed by Moriarity (1975), which Moriarity claims has five advantages:

1. Use of this method will insure that a division will never have the incentive to contract for outside services if the firm provides similar services.
2. The cost allocation process involves a comparison of the cost of providing joint products with the next best alternative.
3. Every division shares in the savings resulting from the decision incurring the joint cost.

4. Some cost is allocated to every division, giving the scheme a notion of fairness.
5. The allocation gives incentives for managers to find the least cost method of providing the firm's joint input.

For the computation of this allocation scheme the firm must know only the cost of providing each division separately and the joint cost of providing each division as a set. Let $EC(K_i)$ be the cost to the firm of each division purchasing the joint input separately, where N is the number of divisions participating in the sharing of the joint cost. Similarly, let $C(EK_i)$ represent the cost of sharing the joint input among the N divisions. Savings to the firm can be represented by $v(N)$ and $EC(K_i) - C(EK_i)$, where the costs are summed over the N divisions. Moriarity states (p. 792) that the cost savings should be allocated based upon each individual division's costs in relationship to the rest of the firm, or $C(K_i)/EC(K_i)$ (again the summations are over the set of N divisions). Accordingly, the allocation for divisions i (α_i) under this formulation is $\alpha_i = C(K_i) - [C(K_i)/EC(K_i)]v(N)$ $i = 1, \dots, n$. But as HHT point out, Moriarity does not investigate under what conditions the allocation is in the core. HHT show that if a core solution does exist for any specific cost function it must be true that

$$\frac{C(EK_i)}{EC(K_i)} \leq \frac{C(EK_i)}{EC(K_i)} \text{ for all } S \subseteq N,$$

where the term on the left-hand side represents the ratio of cost savings by being in the coalition with all divisions and the term on the right represents the ratio of cost savings from being in the coalition S . For the allocation to be in the core, the right-hand side must be greater than or equal to the left-hand side; otherwise some smaller coalition would always break off to provide the input for itself. HHT even provide a counter example with decreasing marginal and average costs to show that it is possible to find situations where a core solution exists but the Moriarity allocation is not in the core. This counter example and its results show that the lack of a core solution makes the incentives to the firm different from what Moriarity supposed, and also lessens the desirability of using this method (especially since decreasing cost industries are the most likely to incur cost allocation regulations). Other methods of allocating costs, such as the Shapley value, however, do have the desirable property of being in the core. This method is discussed next.

3.4.4 The Shapley Value Approach

Another common method of allocating costs among the firm's profit centers is through the use of the method developed by Shapley (1953). Shapley devised his result under the rubric of game theory, but as will be seen, the

allocation is easy to calculate without the express use of a game. Under this approach, each firm is charged an amount equal to the expected incremental cost that is incurred when the participant (or division of the firm) enters the coalition (or set of divisions using the public input), where the order in which the various divisions join the coalition is uncertain. HHT provide the following example. Suppose there is a cost game with three participants such that for participant 1 its cost (C_1) is

$$C_1 = 1/3C\{K_1\} + 1/6[C\{K_1 + 2K_2\} - C\{K_2\}] + \\ 1/6[C\{K_3 + K_1\} - C\{K_3\}] + 1/3[C\{EK_1\} - C\{K_2 + K_3\}]$$

where $C(\cdot)$ represent the cost functions with the corresponding divisions participating in the coalition (that is, $C\{K_1 + 2K_2\}$ is the cost of providing K when divisions 1 and 2 participate in the sharing arrangement).

The incremental costs of division 1 entering the grand coalition (of all divisions) first, second, and third are weighted by the probabilities of entry. If any division is equally likely to enter as any other, then the probability of being first is $1/3$, and the probability of being second, which can occur two ways, is $1/3$, and the probability of being third is also $1/3$. This formula requires each group to pay the incremental cost of being added to the coalition.

Shapley has proven that this result is in the core, but the allocation of joint or common costs based upon

this method still suffers from a defect. For example, suppose the provider of the input does not know how much each division desires of the input. If each division can get a free ride, it has incentives to misstate its actual demand for the input. The Shapley value gives the firm the pareto efficient allocation after the choice to supply a given level of the input has been made.

To determine the optimal amount of the input and to determine the optimal allocation is a two-stage process. First, the optimal amount of the input must be decided upon and then allocation can be obtained. With the Shapley value, the optimal amount of the input is presupposed. HHT show that the Shapley value will not necessarily be in the core if the determination of the input use is subject to strategies behavior.

As can be seen from the last two allocation methodologies, game theory has a role to play in the discussion of allocating costs among multiple users of a public input. Therefore, it is only proper that some discussion of the approach be undertaken.

3.4.5 Game Theoretic Approach to Allocating Costs

This section outlines a game theoretic approach to allocating costs. First, however, a discussion of the particular framework is needed. The allocation of a public input among either a division of a firm or among firms must be a subsidy-free allocation. (Economists generally frown

on subsidies which, in the absence of any positive externality, can cause overconsumption or overproduction and perhaps too many firms producing the good.) Faulhaber (1975) defines this concept of subsidy-free prices in a special way such that "if the provision of any commodity (or group of commodities) by a multi-commodity enterprise subject to a profit constraint leads to prices for the other commodities no higher than [demanders] would pay for themselves, then the price structure is subsidy-free" (p. 966). This type of price structure insures that the provision of each commodity by the firm is pareto superior to non-provision.

A problem with subsidy-free prices is that they may not be welfare maximizing and, in addition, are not morally superior to other non-subsidy-free prices on the grounds of social justice (Fumas and Whinston, 1982). The subsidy-free prices do nothing more than insure that the production and sale of each commodity make all consumers at least as well off as they otherwise would be. Subsidy-free prices may be a welfare improvement, but are at least a floor, so to speak, on welfare and not a set of prices that maximizes welfare. In other words, subsidy-free prices imply that the cost of any one service in joint production with any number of other services is less than or equal to the cost of producing alone. Subsidy-free prices are those prices which satisfy this "stand alone" test.

Given subsidy-free prices as a constraint, a discussion of a second desirable characteristic known as a sustainability is in order. A set of prices, p , is sustainable if there are prices corresponding to a bundle of outputs, q , such that entry by rival firms is unattractive, demand is satisfied, and revenues cover costs of production. The reason why these additional constraints on the game are important is that the results from the game should not generate false price signals that encourage uneconomic entry by potential competitors. This is especially important in the public utilities field. Regulators attempt to make natural monopolies act like a competitive firm in the sense of zero economic profits. This has the potential of causing multiple core allocations, in contrast to the one core allocation generated by the perfectly competitive market.

In order to examine some of these problems, a simple game has been developed based on the ideas of Faulhaber (1975) and Sharkey (1982a, 1982b). Suppose that there is a set of all users of a public input, K , such that $N = \{1, \dots, n\}$ and that there is $S \subseteq N$. Let $C(S)$ be the cost of serving S users as it is also the characteristic function of the game. Generally, the characteristic function is couched in terms of benefits to the participants. In this case, the characteristic function is related to cost

savings through participation in the game. Certain assumptions must also be stated:

1. There is free entry into the market; that is, anyone may construct and use their own K if they so desire.
2. All firms have access to the same technology.
3. Let p be a vector of prices, $p = p(p_1, \dots, p_n)$, such that $\sum p_i = C(N)$. That is, the total cost of serving N customers is equal to the sum of the prices.
4. Let $C(S) + C(T) \geq C(S \cup T)$, where $S \subseteq T = \{\emptyset\}$, which is the condition for subadditivity of costs.
5. Let $\sum p_i \leq C(S)$ for all $S \subseteq N$. This is the stand alone test for subsidy-free prices.

If Assumptions 4 and 5 are met, then the allocation is a core allocation.¹ Let $n^* = |N|$ be the number of users of the public input when $[C(N)/|N|]$ is minimized, where $|N|$ is the cardinality of the set N . The point reflects the average cost of using the public input, assuming everyone pays the same share. Now suppose entry occurs. First, there must be savings left in the use of public input such that the cost function is still sub-additive:

$$C(S+1) \leq C(S) + C(1).$$

If n increased from n^* to $2n^*$ (as in Figure 3.2), $C(N)/|N|$ increases so that at some point between n^* and $2n^*$, there is incentive for a coalition to form that would provide the public input for a different subset $S' \subseteq N$, where $S' \subseteq S$ may not be an empty set. For example, say one firm entered and there are now $n^* + 1$ firms. The entering firm still benefits because $C(N)/(n^* + 1) < C(1)$, but all the other firms, $n < n^*$, now must pay a higher cost

$C(N)/(n^* + 1) > C(N)/n^*$, and so there is incentive for a coalition to form that would lower costs again to $C(N)/n^*$, and one firm would be left out. This creates an unstable cost imputation which does not exist in the core.²

Another result that has not been discussed before concerns the role of economies of scope; the question arises when allocating among different entities--some with scope economies and some without. Suppose, for example, two types of firms existed that used the public input. One firm used the public input to produce one output. So, $C^1(q_{11}(K), q_{12}(K))$ and $C^2(q_{22}(K))$ are the cost functions for these two types of firms, where K is the public input and $C^i(q_{ij}(K))$ is the cost function for firm i with production in market j , for $i, j = 1, 2$.

Some interesting results can be derived if additional restrictions are placed on the cost function for firm 1. Suppose C^1 is subadditive, has declining ray average costs, and is transray convex. These are sufficient conditions for economies of scope.³ In general terms, economies of scope are defined to be savings accruing to a firm that is producing two outputs together as compared to producing them separately, or more formally (for the one firm example)

$$C(q_1(K), q_2(K)) \leq C(q_1(K)) + C(q_2(K)).$$

Suppose also that the second firm decides to join together in the use of K . For both firms the cost of K is

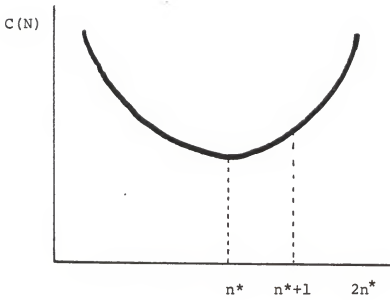


Figure 3.2
Membership Size

lower if sharing is permitted and K is used jointly:

$$C^1(q_{11}(K), q_{12}(K)) + C^2(q_{22}(K)) >$$

$$C^*(q_{11}(K), q_{12}(K) + q_{22}(K)),$$

where C^1 and C^2 are the respective cost functions for firms 1 and 2 producing separately and C^* is the total cost of producing for both firms while sharing K. Thus, there is incentive to form a coalition. Given economies of scope and subadditivity of costs, there exists two types of cost savings: the first from the joint production of good 1 from input K, and the second from the joint use of K by firms 1 and 2. Should firm 2 capture all the savings from the economies of scope and share in the savings due to subadditivity, or should firm 2 be able to share in all savings from the use of K? It seems that if it is possible to do so, the multiproduct firm's benefit flowing to firm should be the result of sharing the use of K. If, on the other hand, the separation of these benefits is not possible, the multiproduct firm will not share with the single output firm if the benefits of economies of scope are larger than the savings generated by forming a coalition to use K:

$$C^1(q_1) + C^1(q_2) = C^*(q_1, q_2) >$$

$$C^1(q_1, q_2) + C^2(q_2) - C^*(q_{11}, q_{12} + q_{22}),$$

where C^* is the cost when both firms are joined in a coalition. If this is true, no coalition will form and the core solution contains one firm. This is a problem

which the Bell System faced prior to the divestiture. AT&T owned and operated the local network and the toll network. If economies of scope were present in the use of the local network for local and toll service, there was a cost savings attributable to this integrated use. Now suppose another company, say MCI, decides to serve a particular locale. Its decision to enter could depend upon the savings attributable to joining with AT&T and the Bell Operating Company in the use of the local network.

The question then arises that if MCI enters, should it be able to enjoy the benefits of economies of scope generated by the AT&T-BOC provision of local and toll service. And if MCI gets that benefit (in terms of cost savings), should it be able to turn around and compete against AT&T in long distance markets. But for the AT&T-BOC economies of scope, MCI may not have entered. AT&T has little incentive to let MCI in to compete and even less when MCI can benefit from AT&T's economies of scope.

As can be seen by the game theoretic approach, there is still not a satisfactory definition of an allocation mechanism although the theory can define a satisfactory result (that is, an allocation in the core). The last section below deals with an approach that is consistent with the game theoretic analysis but adds additional constraints in the form of desired properties that a cost allocation methodology should possess. Using all these

properties, or axioms, to define a cost allocation mechanism helps the situation to a degree, but, as will be shown below, still lacks the results which would solve the allocation problem once and for all.

3.5 Axiomatic Approaches to Allocating Costs

The axiomatic approach to allocating costs actually is consistent with some of the results found in the accounting literature and some of the game theoretic results in that the axiomatic approach lists certain desirable properties that a cost allocation must have in order to meet goals of equity, fairness, and, perhaps to some degree, efficiency. These desirable properties can be taken as constraints to a maximization problem or as rules to a game.

Heany (1982) describes an early axiomatic approach to cost allocation developed by the Tennessee Valley Authority in the 1930s. The TVA's problem was how to allocate the cost of a reservoir among flood control, navigation, fertilizer production, national defense, and power generation. Three criteria (or rules) were developed which the cost allocation scheme should satisfy to be a "good" mechanism:

1. No individual should be charged more than his alternative cost or maximum benefit.
2. No subgroups should be charged more than its alternative cost or maximum benefit.
3. Each participant should share in the savings.

The first criterion permits a cost allocation to an individual which is between the cost of obtaining the service from some other means and the maximum amount of benefit the individual obtains from the service. This is an attempt to give incentives to join in the coalition to provide service. The second criterion is the bounds on the cost allocation for a given group. A cost allocation mechanism is sometimes referred to as individually rational and group rational if it satisfies the first two criterion. The third criterion is an attempt at implementing fairness. Everyone should receive some net benefits from participating in the sharing of these costs.

Loehman and Whinston (LW) (1974) attempt to set up a set of axioms, like the TVA's, which will lead to an allocation system which is "fair." LW propose their own set, consisting of the following axioms:

1. Charges for the use of the facility must cover costs.
2. A user's charges are to be based only on the incremental costs caused by the user and not the incremental costs of other users.
3. The charge to a user is independent of labeling or ordering of users. Those that demand the same amounts and cause the same incremental cost should pay the same charge.
4. The charge is homogeneous of degree one in the incremental costs. If all incremental costs double, then the charge will double.

Axiom (1) simply requires that the firm break even and (2) employs the notion of incremental costs for efficiency purposes. Axiom (3) is a principle of horizontal equity,

and (4) requires "proportional" equity (that is, big users pay more than small users).

LW (p. 246) prove that, using this set of axioms that given n users with fixed positive demands for the public input K , the charge for the use of the input will be

$$\alpha_i = \sum_{S \subset N} \left[\frac{(n-s)! (s-1)}{n!} \right] [C(\sum_S K) - C(\sum_S K - \{i\})],$$

where s is the number of users belonging to the coalition. This is just the Shapley value for $S \subseteq N$, where the summation over all participants in the coalition of the probability of entry multiplied by the incremental costs yields the allocation to user i . So, this list of axioms results in a cost allocation mechanism which exists and, in fact, is one that has been shown to exist in the core.

Before continuing, it is important to note that the idea behind axiomatic cost allocation schemes is finding a set of rules which are desirable and agreeable to all concerned. The rules are then applied to determine how to share the benefits of sharing an economic indivisibility resulting from economies of scale, or joint costs. These allocations, like subsidy-free prices, are not necessarily welfare maximizing prices or allocations--they yield only a given level of welfare. The benefit to employing either an axiomatic scheme or one that employs the notion found in subsidy-free prices is that society is guaranteed a level of welfare, while under some other scheme the welfare could

be higher, but could be much lower. By using these schemes, society is able to avoid huge welfare losses.

More powerful axioms have been developed to deal explicitly with the problem increasing returns to scale. Billera, Heath, and Raanan (1978) used a Aumann-Shapley (AS) value for non-atomic games (where the players range over a continuum) to allocate costs for a common telephone system to the various departments at Cornell University. In a non-atomic game the player space is a continuum rather than, say, outputs. For example, in the allocation of costs for Cornell University's telephone system, the player space is the time spent on various types of phone calls. In atomic games, the player space can be landing rights, as an airport (Littlechild, 1972), or perhaps access to a telephone system. The conceptual difference centers around the notion of whether the players are discrete units or continuous units. Thus, the major benefit of AS allocations is that it can allocate something that is infinitely divisible.

The difficulty, however, with the AS result is that the axioms upon which the mechanisms are based are only applicable to the case where there is no fixed costs (even though there are scale economies). As Miriman, Samet, and Tauman (MST) (1983) point out, this is not a major problem in the long run since, by definition, there are no fixed costs in the long run. But, the optimal short-run tech-

nology may, and often does, have a fixed cost component which must be allocated among its outputs if regulators so require.

MST examine the problems of the AS allocator and derive an allocation mechanism based on the AS allocation that allows the allocation of fixed costs for the short run firm. AS developed the following five axioms:

1. Cost sharing. The sum of the allocations determined by the mechanism must cover the cost of provision.
2. Rescaling. If the scale of measurement changes for the commodities, the allocation is changed accordingly.
3. Consistency. Each unit of the "same good" has the same price.
4. Additivity. Whenever a cost function can be split into two or more components the sum of the allocations to a particular division is the same as the allocation would be if the cost function was not broken into components.
5. Positivity. Given two different cost functions, where $C1 < C2$, the allocation derived from $C2$ will be no lower than that derived from $C1$.

Axiom (1) is similar to the breakeven axiom set up by LW and the Moriarity equity provision. Axiom (2) is merely LW's homogeneity axiom. Axiom (3) can be restated as the proposition that, all other things held constant, red cars should be priced the same as blue cars given that red paint costs the same as blue paint. Since there are no demand conditions in the axioms, Axiom (3) must be true. Ramsey prices, which are an example of demand compatible cost allocation mechanisms, account for these demand differences to allocate costs by using the demand elasticity in the

allocation formula. In the absence of demand information, Axiom (3) insures a "sort" of horizontal equity.

Axiom (4) is used to produce the result that if costs increase the allocation to a particular product should increase, and Axiom (5) requires that the allocations from separable parts of the cost function (for example, production costs and management costs) sum to the allocation if the costs of management and production were not separated.

With these five axioms it has been proven (Billera and Heath) that, given continuous cost functions with no fixed cost components, there exists a unique welfare allocation which is the AS allocation.

Now, let $C(q^*) = V(q^*) + K$, where q^* is the cost minimizing vector of outputs, C is total costs, V is variable costs, and K is the fixed costs or public input. This, of course, is a short-run cost function. However, using AS prices to allocate $C(q^*)$, MST claim that an allocation of fixed costs is associated with the allocation of total costs when the technology used is the optimal technology. Since the AS prices cover both fixed and variable costs, it is possible to "break down" the allocation to variable and fixed costs. If $C(q^*)$ is used to determine the AS cost sharing prices, there will be a resulting set of prices $p \equiv p(p_1, \dots, p_m)$ when q^* is produced such that:

$$p_i = \int_0^1 \frac{\delta C(tq)dt}{\delta q_i} \quad \text{for } i = 1, \dots, m,$$

where t is a scalar. Since prices cover both variable and fixed costs, it is possible to calculate another set of prices due only to the variable cost function, so

$$v_i = \int_0^1 \frac{\delta V(tq)dt}{\delta q_i} \quad \text{for } i = 1, \dots, m.$$

Therefore, the difference in prices is

$$\sum_m^i (p_i - v_i)q_i = C(q^*) - V(q^*) = K,$$

which means $(p_i - v_i)q_i$ is the part of the price, p_i , which covers K for output of division i . It follows that

$$\alpha_i K = (p_i - v_i)q_i \Rightarrow$$

$$\alpha_i = (p_i - v_i)q_i / K \Rightarrow$$

$$\sum \alpha_i = 1.$$

So, with just a "break down" of costs between the fixed and variable components, it is impossible to discover the implicit allocation of fixed costs which the Aumann-Shapley value provides.

Nevertheless, this axiomatic approach has some useful results, the most important of which is that this list of rules derives a unique allocation which exists in the core. The problem, however, is that this mechanism allocates variable costs as well as fixed costs. What if there were variable costs which were directly allocatable to a particular division or user. Under AS, these costs

would be allocated among all users even though only one division caused them to be incurred.

In addition, it is not clear that this allocation would be sustainable against competition since there is no demand side considerations in the mechanism. Demand is, for all intents and purposes, completely inelastic. This is not a valid assumption for an implementable allocation scheme. Finally, although related to the sustainability issue but important in its own right, is the effect upon entry. As a general result, AS would protect rivals from predatory prices as defined by Areeda and Turner (1975), or Ordover and Willig (1979), but the AS allocation could actually make entry less desirable for some markets and more desirable for others. It would be interesting to see the result of an AS allocation when additional axioms concerning entry and sustainability of the firm are added.

3.6 Conclusion

This chapter has shown five types of cost allocating mechanisms. The first, joint cost-Lindahl approach, was shown to be the efficient allocation mechanism. The second approach, Ramsey pricing, is the benchmark for performance for other allocation rules. In the absence of the efficient allocation due to free-riders or the principal-agent problem for the firm, Ramsey prices represent the constrained welfare maximizing result. Given Ramsey prices as a standard, other allocation mechanisms based

upon the idea of fully distributed costs fall short of the efficiencies guaranteed by the Ramsey rule. The game theoretic approach yielded information about the end result of the allocation mechanism. A desirable property is that the allocation involves a core solution. This implies that there is at least a pareto efficient solution and helps to eliminate allocation mechanism from consideration. Finally, axiomatic schemes were discussed to show that by obtaining a list of agreeable axioms it was possible to get a core solution for the allocation. The problem with the axiomatic approach is that it did not cover all the possible axioms needed for many cases. Axioms concerning sustainability, demand compatibility, and directly allocatable costs must be added to see if desirable allocation still exists.

Cost allocation is a very important part of today's administrative regulation and will become even more important in the future (to the dismay of many positive economists). The access charge arrangements currently being implemented in the telecommunications industry are, at best, a conglomeration of the above approaches, or at worst, a completely ad hoc method. The business of allocating costs has a long history, but the problems of railroad regulators from the latter 19th century are still with 20th century regulators. Hopefully, these same problems will not be around for another hundred years.

3.7 End Notes

1. See Fumas and Whinston (1982) for a fuller discussion of this point. Pareto efficient outcomes are not comparable on a basis of fairness.

2. Proof:

This proof requires two steps: (1) feasibility and (2) nonimprovability.

1. Feasibility is rather trivially shown by

$$\sum p_i < C(N).$$

2. Nonimprovability. Say $\sum p_i > C(S)$. Then, if there exists some S' in S such that $\sum p_i < C(S')$, then there is incentive to form another coalition S' for these in S' , so $\sum p_i > C(S)$ is not in the core.

3. This problem is somewhat similar to the results described by Charles Tiebout (1956), especially in the characterization of the "average cost" of public goods provision. Even though it is intuitively similar to the average cost function of a firm, it has a distinct difference: the Tiebout average cost is a characteristic of demand only while regular average cost terminology refers to a supply condition.

While on this subject, the whole area of club theory is particularly relevant to the problem presented in the text of the Chapter II, especially for multiproduct clubs. Sandler and Tschirhart (ST) (1980) talk about this phenomena. ST specifically mention a paper by Berglas and Pines (1978) that develops a model for a multiproduct club which provides a better theoretic foundation (according to ST) for public goods analysis. The problem with the Berglas and Pines article, though, is that two important parts of the models in the text were not included: cross subsidization and complementary in production (economies of scope).

Finally, it should be noted that the literature surrounding the sustainability analysis by Baumol et al. (1982) of multiproduct firms is very relevant to the concerns of a multiproduct club. ST point out, however, that to do this right requires that the sharing group size for each product must be introduced to the cost functions, and the particular indivisibility of the multiple products

4. See Baumol (1977) for a complete proof of this assertion and Bailey and Friedlander (1978) for an overview of the theory of multiproduct production.

CHAPTER IV COST ALLOCATIONS WITHIN A MULTIPRODUCT FIRM

The types of questions arising from a regulation determined cost allocation mechanism are different from those concerning the optimal allocation. Regulators use the allocation as a method of achieving policy goals. Over the last forty years, the FCC and the States have altered the cost allocation mechanism numerous times. An interesting question to examine is what happens to the incentives facing the firm when the allocation mechanism is changed. This chapter examines the interrelationship between the cost allocation mechanism and the demand for the capital input when the input is shared by two fixed-proportion divisions of the firm. It is useful to examine the fixed-proportions case to isolate the interrelationships before examining the more general case which includes the added complexity of substitution between capital and other inputs.

4.1 Introduction: Production with One Input

The cost allocation problem can be seen in a model using fixed proportion production functions, but before continuing, an even simpler model is employed to set up the analysis. Assume that in each of two markets the production process uses only capital and takes on the following form:

$$q_1 = K \quad (4.1.a)$$

$$q_2 = K, \quad (4.1.1b)$$

where K represents the common capital used in the production of goods 1 and 2. Certain assumptions about this K need to be stated. First, once K is selected by the firm it is fixed in the short run. Second, once the divisions have agreed to jointly purchase a given level of K , a change in the demand for K by either division does not affect the use of K by the other division. In the public goods literature, this property is known as nonrivalry in consumption--it implies that use of the capital by one division does not effect the other division's use. We also assume that there is no congestion in the use of the input. Finally, users of K can be excluded from use by a positive price. The second two properties imply that K is a price excludable public good while the first property will allow analysis of the short-run welfare effects of a change in the allocation mechanism.

Given the two production functions, the objective is to maximize profits subject to a regulatory constraint that requires the firm to earn zero economic profits. The problem then becomes:

$$\max \pi = R^1(q^1) + R^2(q^2) - rK \quad (4.1.2)$$

subject to

$$R^1(q_1) - \alpha rK = 0 \quad (4.1.3)$$

$$R^2(q_2) - (1 - \alpha)rK = 0 \quad (4.1.4)$$

$$q_i(p_i) = q_i(K), \quad i = 1, 2; \quad (4.1.5)$$

where $R^i(q_i)$ represents the revenue function for good i , r represents the input wage rate for K , and α is the cost allocation factor which is assumed to be between 0 and 1. In addition, it is assumed that both markets, demand for the joint input K presses on capacity such that $q_i(p_i) - q_i(K) = 0$. The output q_i is assumed to be strictly positive in order to insure that the firm's decision-making process yields a nonzero solution. From equations (4.1.1a) and (4.1.1b) it can be seen that $q_1 = q_2 = K$. By substituting (4.1.1a) and (4.1.1b) into (4.1.3) and (4.1.4), the constraints become:

$$R^1(q_1) - \alpha r q_1 = 0 \quad (4.1.3a)$$

$$R^2(q_2) - (1 - \alpha) r q_2 = 0 \quad (4.1.4a)$$

Since this next analysis is a short-run analysis, once the firm selects the level of K , the firm cannot alter its choice because capital is fixed. The firm has one chance to pick its level of capital. Equations (4.1.3) and (4.1.4) imply that price equals average cost in both markets where α represents the share of total capital costs allocated to product 1 and $(1 - \alpha)$ represents the share of capital allocated to good 2. For simplicity, assume the following linear demand curves exist for the two goods:

$$P_1 = A - \tau_1 q_1 \quad (4.1.5)$$

$$P_2 = A - \tau_2 q_2. \quad (4.1.6)$$

This gives a system of six equations (4.1.1a, 4.1.1b, and 4.1.3-6) and five unknowns (p_1 , p_2 , q_1 , q_2 , K) for the system. To solve the system, it is necessary to determine α , the cost allocation mechanism. Solving equations (4.1.3) and (4.1.4) for K yields

$$\frac{A - \alpha r}{\tau_1} = K = q_1 \quad (4.1.3b)$$

$$\frac{A - (1 - \alpha)r}{\tau_2} = K = q_2 \quad (4.1.4b)$$

and the only way that (4.1.3b) and (4.1.4b) can both be true is if they are equivalent. Equating (4.1.3b) and (4.1.4b) and solving for α yields

$$\alpha = \frac{A(\tau_2 - \tau_1) + \tau_1 r}{r(\tau_1 + \tau_2)} \quad (4.1.7)$$

If it is assumed that the demands for goods 1 and 2 are identical ($\tau_1 = \tau_2$), then equation (4.1.7) yields $\alpha = (1/2)r$. So, given identical demands and an identical production process, the optimal allocation of costs occurs when both markets are charged one-half of the total capital costs.

If the demands are not identical, a necessary question to ask is what happens to α . If, for example, $\tau_1 > \tau_2$ then $\alpha < (1/2)r$. In Figure 4.1, the two demand curves are drawn reflecting $\tau_1 > \tau_2$. It can be seen that the market with the greater relative demand pays more towards covering the cost of capital. The Ramsey rule where price is selected

based upon the elasticity of demand is consistent with the result shown here. Let marginal cost for market 1 be αr and for market 2, $(1 - \alpha)r$. In Chapter III it was shown that the Ramsey rule required that

$$\frac{(p_1 - \alpha r)e_1}{p_1} = \frac{(p_2 - (1-\alpha)r)e_2}{p_2} = v,$$

or that the percentage deviation from marginal cost multiplied by the output demand elasticity across all demanders is equal to a constant, v . In this example, where α is chosen in the manner above, $v = 0$. This implies that a welfare maximizing α is chosen. Each division pays only the marginal valuation for the use of capital.

The pricing mechanism selected to make the demand for K identical by both production processes is called a Lindahl (1959) price mechanism. Instead of allowing price to vary and selecting the corresponding quantity of output as is done in private goods analysis, a Lindahl price is the price system that, given a level of a public good (K), yields an individualized price for each user of the public good where the sum of the individual prices is equal to the marginal cost of providing the public good. Not only is αr the marginal cost for division 1, but it is also the Lindahl price for division 1, and $(1 - \alpha)r$ is the Lindahl price for division 2. In addition, the cost allocation mechanism picked in this manner yields the Samuelson-

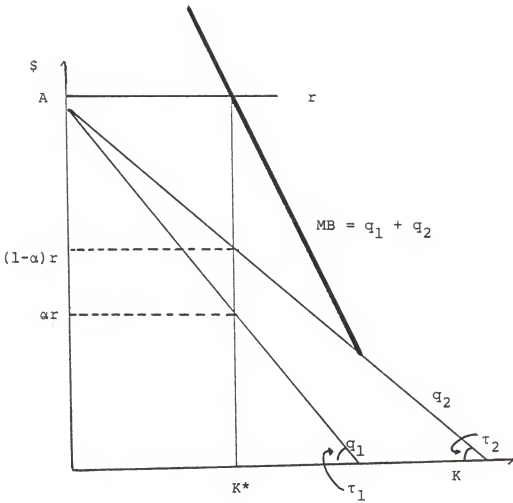


Figure 4.1
Optimal Provision of K

ian conditions (Samuelson, 1953) for the optimal provision of a public good where the sum of the marginal evaluations of capital for each division is equated to the marginal cost. In Figure 4.1, the optimal provision is described graphically. A vertical summation of the demand curves for each division ($q_1 + q_2$) yields the total marginal evaluation for capital (Σq_i) curve for the joint use of K . Since $K = q_1 = q_2$ the optimal cost allocation scheme is based upon the marginal evaluation for K which yields the Lindahl price of αr for division 1 and $(1 - \alpha)r$ for division 2. From this simple model, it can be seen that a firm will achieve an efficient result if it is allowed to choose the allocation between markets given that it must price at average cost in both markets and produce a strictly positive level of output.

4.2 Fixed Proportions Model

To make the analysis a bit more realistic, two assumptions are modified. First, both demands are distinct, so equations (4.1.5) and (4.1.6) are replaced by

$$p_1 = A - \tau_1 q_1 \quad (4.1.5a)$$

$$p_2 = B - \tau_2 q_2. \quad (4.1.6a)$$

And second, the variable input z_i is used in the production of q_i , so that

$$q_1 = \min [\beta_1 z_1, K] \quad (4.1.1a)$$

$$q_2 = \min [\beta_2 z_2, K] \quad (4.1.1b)$$

Similarly, equations (4.1.2), (4.1.3), and (4.1.4) are modified so that the problem is

$$\max \pi = R^1(q_1) + R^2(q_2) - wz_1 - wz_2 - rK \quad (4.1.2a)$$

$$\text{s.t. } R^1(q_1) - wz_1 - \alpha rK = 0 \quad (4.1.3a)$$

$$R^2(q_2) - wz_2 - (1-\alpha)rK = 0. \quad (4.1.4a)$$

Once again, to solve for α , the two regulatory constraints from (4.1.3a) and (4.1.4a) must first be solved for K , which yields

$$K = \frac{A - w_1 - \alpha r}{\tau_1} \quad (4.1.3b)$$

$$K = \frac{B - w_2 - (1-\alpha)r}{\tau_2} \quad (4.1.3b)$$

where $w_i = w_i/\beta_i$. Equating (4.1.3b) and (4.1.4b) and solving for α yields

$$\alpha = \frac{\tau_2(A - w_1) - \tau_1(B - w_2) + \tau_1 r}{r(\tau_1 + \tau_2)}.$$

This methodology yields αr and $(1 - \alpha)r$ which are the Lindahl prices for divisions' 1 and 2 joint use of K . This is the α which maximizes welfare given the average cost price constraint in both markets. The next step in the analysis is to examine the effect of changing on the conditional input demands for the firm. Solving Equation (4.1.3) for z_1 , Equation (4.1.4) for z_2 , and Equations (4.1.3) and (4.1.4) jointly for K yields

$$z_1 = \frac{A - w_1 - \alpha r \beta_1}{\tau_1 \beta_1^2} \quad (4.1.8)$$

$$z_2 = \frac{B - w_2 - (1 - \alpha)r\beta_2}{\tau_2\beta_2^2} \quad (4.1.9)$$

$$K = \frac{(A - w_1) + (B - w_2) - r\tau_1}{\tau_1 + \frac{1}{2}} \quad (4.1.10)$$

If the government decided for some policy reason to alter α to allow low income consumers to be able to afford product 1, how would the firm's decisions be altered? By taking the derivative of the input demand with respect to α , the change in input demand can be seen.

$$\frac{dz_1}{d\alpha} = -\frac{r}{\beta_1\tau_1} < 0 \quad (4.1.11)$$

$$\frac{dz_2}{d\alpha} = \frac{r}{\beta_2\tau_2} > 0 \quad (4.1.12)$$

Equation (4.1.11) shows that if α increases, the amount of z_1 demanded falls since all terms are positive. This makes intuitive sense. As capital becomes more expensive in market 1 due to an increase in α , the firm must continue to price at average cost and earn zero economic profits in both markets. The only way to do this is to reduce the amount of z_1 employed. In Figure 4.2, the division starts out at the optimal point denoted A. When the cost allocation factor changes, the firm experiences a reduction in the amount of capital needed to produce with a new level of z . The amount of K , however,

is insensitive to the change in α since K is fixed in the short run (equation 4.1.13).

Point B in Figure 4.2 reflects the point the division would desire to move to if it could lay off the excess capital. The division, however, must pay its share of the firm's entire capital stock. This prohibits the firm from operating at point B. Given an increase in α and the fact that it must pay a share of the entire costs of K^* , the division must lower output and lay off labor to a point such as C. At point C, the firm is paying for K^* amount of capital, but because the capital is so expensive relative to the division's original position at point A, the firm lays off its variable input and "wastes" the capital input. Waste in this sense means that the division would not normally choose the level of K it is now paying for if the division was still operating efficiently.

As a result of the increase in α , output falls from q_1^* to q_1^0 , which is a greater loss of output than if the firm could lay off some of K . Equation (4.1.12) shows the effect on the variable input in market 2. As α increases, the amount of capital charged to the second division decreases. Since K is fixed (equation 4.1.13), the firm cannot expand output and since the division is earning rents, it is violating the regulatory constraints. If the firm were to lower prices in division 2 to dissipate the rents, it would cause an increase in quantity demanded

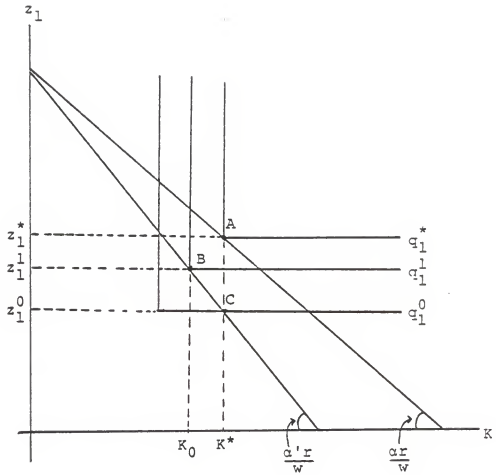


Figure 4.2
Effect of a Change in α in Market 1

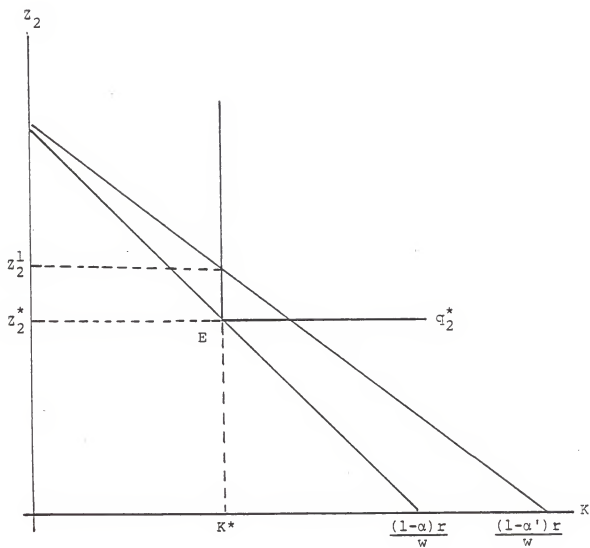


Figure 4.3
Effect of a Change in α in Market 2

which, in turn, would be impossible to fulfill given the additional constraint that all demanders must be served. Since the firm cannot increase K or lower price, it can only use more variable inputs to "soak up" the rents accruing to the firm in case of an increase in α .

Figure 4.3 shows how this occurs. As α increases, $(1 - \alpha)$ decreases which rotates the isocost line away from the origin. The firm cannot move to a higher production isoquant because of the capital is fixed at K^* and the fixed proportions production function. Division 2 hires more z_2 to allow it to dissipate profits.

Figure 4.4 shows the effects of a changing α on the two divisions of the firm. Starting at price equivalent to average cost given α , each division selects the same level of K which is the optimal K^* . Each division then produces output based upon K^* such that $q_1 = q_1(K^*)$ and $q_2 = q_2(K^*)$. If α is increased to α' then the average cost in division 1 becomes $(w + \alpha'r)$ and in division 2 average cost falls to $(w + (1 - \alpha')r)$. It seems that in division 1, the response of a change in α would be to produce at q_1 (point B in Figure 4.4). This is not correct since the division must still pay for a share of K^* amount of capital and must produce at q_1^0 (point C in Figure 4.4). In fact, what actually occurs is an even larger decline in output as α increases. Output falls from $q_1(K^*)$ to q_1 because division one's wage bill for K is $\alpha'K^*$. In order

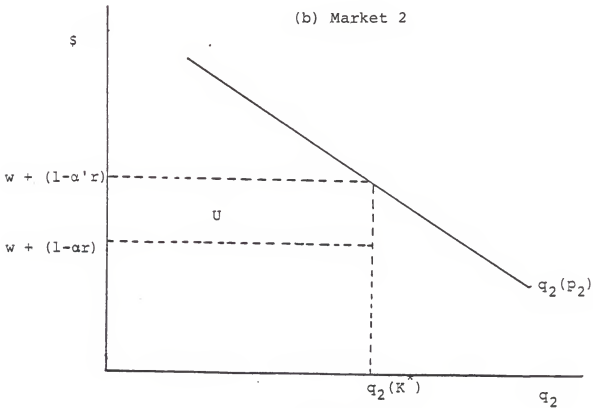
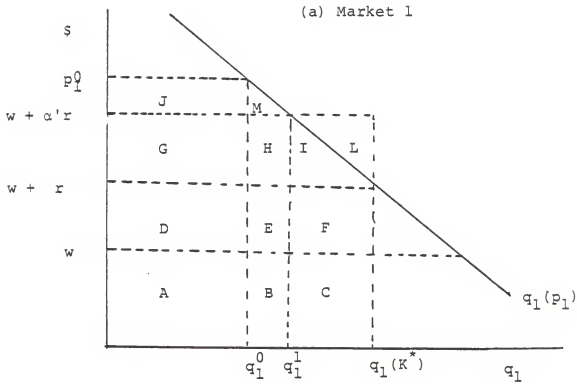


Figure 4.4
Welfare Effects

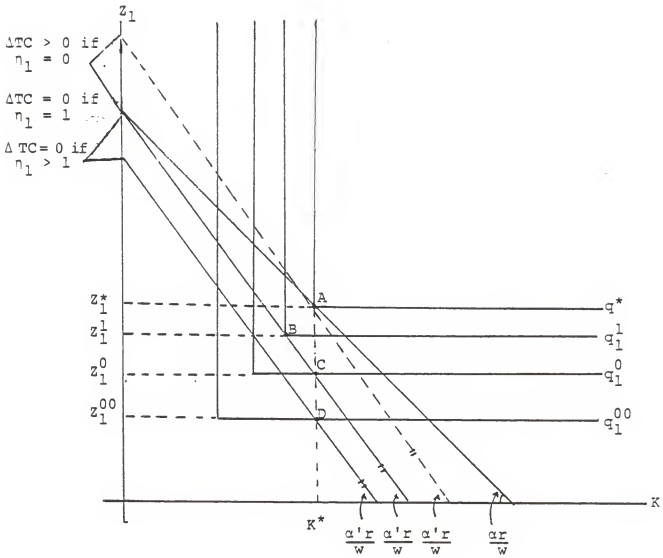


Figure 4.5
Changes in Market 1 under Different Assumptions
About Demand Elasticity

to pay for the total amount of allocated capital, the firm must raise price so that it breaks even. Price thus rises from $(w + \alpha'r)$ to p_1^0 to ensure that total revenue equals total cost. Figure 4.4b shows what happens in market 2 as average cost falls. The division experiences a cost savings but cannot increase output past $q_2(K^*)$ because of the fixed nature of the investment.

If we assume that there are zero income effects, we can see that the net effect in terms of efficiency is negative. As output falls from $q_1^*(K^*)$ to q_1^0 due to an increase in α it is possible to show that even if all the cost savings in division 2 is recognized as a social cost saving, the loss to society will be larger than the dead-weight loss (MHI) due to a loss in output in division 1. In division 1 there is no change in profits, as they still are zero. The loss to consumers, however, is GJMHI. To show the loss in efficiency, it is necessary to show that the cost saving in division 2 is less than GJHMI. If we assume the elasticity of demand in the relevant range is unitary, then examining Figure 4.4,

(a) $GJ = BEFC$;

(b) since total costs in division 1 do not change, the capital cost increase must equal the labor cost decrease (BC);

(c) since the cost savings from division 2 equals the cost increase to division one, $U = BC$;

By combining (a), (b), and (c), the cost savings in division 2 equals BC: the reduction in variable inputs in

division 1. The area BC, however, is less than BCEF = GJ. So the social loss exceeds the area MHI by the area EF. So even if the cost increase in division 1 is exactly offset by the cost savings in division 2, welfare still falls.

By altering the assumption of unitary elasticity of demand for the output in market 1, more information can be gleaned from this model due to output changes that occur as a result of a change in α . Equations (4.1.11) and (4.1.12) can be rewritten so that

$$\begin{aligned} z_{1\alpha} &= -r/\beta_1\tau_1 = - (rq_1/\beta_1p_1)e_1 &< 0 \\ z_{2\alpha} &= r/\beta_2\tau_2 = (rq_2/\beta_2p_2)e_2 &> 0 \end{aligned}$$

by noting that $\tau_i = dp_i/dq_i$ and by multiplying the expression by $(q_i/p_i)(p_i/q_i)$. This allows us to obtain the derivative of z_i with respect to α to be a function of the output demand elasticity (e_i). If e_i , for example, is zero, then there will be no change in the amount of z_i demanded by division 1. If, however, e_i is perfectly elastic, an impossibility result occurs. An attempt to make division 1 pay any more of the joint costs causes the firm to shut down division 1. The more elastic the output elasticity (all other things held constant) for division 1, the greater will be the reduction in z_1 . This is shown diagrammatically in Figure 4.5.

When the isocost curve is allowed to change as output changes, there is a scale effect or output effect. Output may change when input prices change. For example, the

own price elasticity is zero, there is no change in output and the division will remain at point A in Figure 4.5 even after a change in α . In order to remain at point A, however, the division must spend more on inputs and total cost will rise. The optimal amounts of inputs will be used when the own price demand elasticity is perfectly inelastic. As the demand elasticity becomes more elastic, however, output does change and the change in total costs becomes smaller. So the smaller the demand elasticities, the smaller will be the distortions due to a change in α . Figure 4.5 shows the same result as Figure 4.2 when the elasticity of demand is unitary. Production falls to q_1^1 and z_1 falls to z_1^1 . As the output elasticity becomes greater than 1, the isocost curve shifts down towards the origin and the firm will lay off more z_1 to produce, using K^* as shown by point D. Total cost falls with higher elasticities and the fall is solely due to a decrease in the variable input since the division must pay for the entire amount of capital (K^*) even if it is not used.

4.3 Long Run Analysis

The fixed-proportions firm with constant returns to scale technology, sharing of a capital input, and a linear demand for output, alters its long run use of capital based upon the relative sizes of the output demands. Recall equations (4.1.3b) and (4.1.4b) which together determine the joint demand for K . By taking the total derivative of

each division's demand for the capital input it is possible to see the effects of a altered α on the long run choice of K .

$$dK/d\alpha = -r/\tau_1 < 0; \text{ for division 1,}$$

$$dK/d\alpha = r/\tau_2 > 0; \text{ for division 2.}$$

Since the overall change in K depends upon the effect of a change in the joint use,

$$dK/d\alpha = r(1/\tau_2 - 1/\tau_1) \gtrless 0. \quad (4.3.1)$$

As the difference in the inverses of the respective slopes gets larger, the difference in the long run amount of capital increases. In addition, if the slopes are equal there will be no change in the long run demand for capital.

This is exactly what one would expect. It is important to note that if the cost allocation mechanism is altered then there can be a change in the demand for capital. By altering the allocation mechanism, there is an effect on the firms incentives to chose the long run capital stock. This same effect will be explored later in Chapter Five for the case of variable proportion production functions, but in both cases there is the same general effect as the allocator is altered. This is something that policy makers never realized when they increased the allocator in the telecommunications industry.

The effect of a new allocator is not necessarily neutral, and may in fact cause total firm costs to rise in the long run. A reallocation of costs from (say) local

service to (say) long distance could cause the size of the local network to diminish. This could have the effect of effectively increasing the cost of local service in terms of higher prices, or may impose costs in terms of lower efficiency on toll users.

4.4 Increasing Returns to Scale in One Market

A logical question that follows from this type of analysis is what are the firm's incentives if there are increasing returns to scale in one market. This analysis requires a more rigorous formulation. This change in the model's formulation will allow some assessment to the current state of affairs in the telecommunications industry. If one of the markets using a joint input consistently gets more and more of the joint costs allocated to it, the types of distortions occurring vis a vis the output mix and capital input mix need to be analyzed. Assume that division one's production function is now

$$q_1 = \min [\beta_1 z_1, K^2].$$

The problem, then, is to maximize profits subject to the break-even provision and the fixed proportions requirements. This becomes

$$\begin{aligned} \max \pi &= p_1(q_1)q_1 + p_2(q_2)q_2 - wz_1 - wz_2 - rK \\ \text{s.t. } p_1 &\leq (wz_1 + \alpha rK)/q_1 \\ p_2 &\leq (wz_2 + (1 - \alpha)rK)/q_2 \end{aligned}$$

$$K^2/\beta_1 = z_1$$

$$K/\beta_2 = z_2$$

$$q_1 = \min [\beta_1 z_1, K^2]$$

$$q_2 = \min [\beta_2 z_2, K].$$

By substituting K^2/β_1 for z_1 and K/β_2 for z_2 , the problem can be written as:

$$\begin{aligned} \max L = & p_1(q_1)q_1 + p_2(q_2)q_2 - w/\beta_1 K^2 - w/\beta_2 - rK \\ & + \mu_1[w/\beta_1 + \alpha r/K^2 - p_1] \\ & + \mu_2[w/\beta_2 + (1 - \alpha)r - p_2], \end{aligned}$$

which yields the following first order conditions,

$$\begin{aligned} L_K = & (1 - \mu_1)(MRP_K^1 - 2wK/\beta_1) + (1 - \mu_2)(MRP_K^2 - w/\beta_2) - r \\ & + \mu_1 \alpha r + \mu_2(1 - \alpha)r \leq 0 \end{aligned}$$

(4.3.1)

$$L_K K = 0 \quad (4.3.1a)$$

$$L_\alpha = rK(\mu_1 - \mu_2) \leq 0 \quad (4.3.2)$$

$$L_\alpha \alpha = 0 \quad (4.3.2a)$$

$$L_{\mu_1} = -p_1(K) + w/\beta_1 + \alpha r/K \geq 0 \quad (4.3.3)$$

$$L_{\mu_1} \mu_1 = 0 \quad (4.3.3a)$$

$$L_{\mu_2} = -p_2(K) + w/\beta_2 + (1 - \alpha)r \geq 0 \quad (4.3.4)$$

$$L_{\mu_2} \mu_2 = 0 \quad (4.3.4a)$$

where MRP_K^i is the marginal revenue product of capital for market i , $i = 1, 2$. By allowing the firm the choice of K and α given the regulatory constraints, the firm will choose an allocator (α) and a capital stock (K) to maximize profits. Given that this is a Kuhn-Tucker problem there are numerous possible solutions, one of which deserves attent-

ion. If $K > 0$, $\alpha > 0$, and $\mu_1 > 0$, the first order conditions can be rewritten as

$$p_1 = w/\beta_1 + \alpha r/K \quad (4.3.5)$$

$$p_2 = w/\beta_2 + (1 - \alpha)r \quad (4.3.6)$$

$$\mu_1 = \mu_2 \quad (4.3.7)$$

$$MRP_K^1 - 2wK/\beta_1 + MRP_K^2 - w/\beta_2 = 0. \quad (4.3.8)$$

Now, since the marginal revenue product equals marginal revenue multiplied by the marginal product we know that at the corner of the isoquant the marginal product for capital for division one is $2K$, and for division two the marginal product is 1. The marginal revenue product of capital for division i can be written as

$$MRP_K^i = MR_1^i * MP_K^i = \left(\frac{\delta p_i}{\delta q_i} q_i + p_i \right) * MP_K^i = p_i (1 - 1/e_i) * MP_K^i,$$

where e_i is the absolute value of the price elasticity of demand. Equation (4.3.1) can be rewritten as

$$p_1(1 - 1/e_1)2K + p_2(1 - 1/e_2) - 2wK/\beta_1 - w/\beta_2 = 0.$$

The next step is to solve for K and then take the total derivative of K with respect to α to see the long run change in K as the allocation mechanism is altered when one division experiences economies of scale. Finding K , we get

$$K = \frac{p_2(1 - 1/e_2) - w/\beta_2}{p_1(1 - 1/e_1)2 - 2w/\beta_1}. \quad (4.3.9)$$

Now by taking the derivative with respect to α , and letting $m_i = (1 - 1/e_i)$, we obtain

$$dK = \{[(\delta p_2 / \delta K) m_2 dK + (\delta p_2 / \delta \alpha) m_2 d\alpha](2p_1 m_1 - 2w/\beta_1)\} * D^{-2} \\ - \{[(\delta p_1 / \delta K) m_1 dK + 2(\delta p_1 / \delta \alpha) m_1 d\alpha](p_2 m_2 - w/\beta_2)\} * D^{-2}$$

where D is the denominator of (4.3.9). Solving

$dK/d\alpha$ we obtain,

$$\frac{dK}{d\alpha} = \frac{-r[m_2 Q_1 + m_1 Q_2 / K]}{2Q_1^2 + \alpha m_1 Q_2 / K^2} \quad (4.3.10)$$

where $Q_i = [MRP_K^i - w/\beta_i]$, and the derivatives of prices were obtained from the constraints in equations (4.3.3) and (4.3.4). Thus an increase in α (to the jurisdiction with economies of scale) can cause the long run K stock to fall or rise.

Notice that the size and the sign of the effect of a change in the allocator depends upon the relative price elasticities embodied in the m_i terms. The overall sign of $dK/d\alpha$ depends on both the sign of the numerator and of the denominator. With the existence of the various sign possibilities, a unique determination of the sign is impossible. Even so, some information can still be gleaned from the derivative to give some information about the direction of the change. Looking first to the denominator, it is likely that the denominator will be positive since (a) the first term will always be positive and (b) the second term (even if negative) will most likely be small since it is multiplied by α (which is less than 1) and is

also divided by K^2 . The term within the brackets in the numerator is also likely to be positive, but there is the possibility that it could be negative.

First, if we assume that the denominator is positive, and that the term within the brackets is also positive, then $dK/da < 0$. This implies that as a is increased, the price increases (because the division experiencing economies of scale and must ride up its "average cost" curve) -- lowering the quantity demanded for market one. This, in turn, lowers the derived demand for K for market one which outweighs any increase in the derived demand for K in market two. This result would be consistent with a relatively inelastic demand in market two and a relatively elastic demand in market one. The benefits from scale economies are thus frustrated by increasing the allocation to the division with scale economies.

Now, if we assume that the relative elasticities are reversed and market one (the market with economies of scale) was the relatively inelastic and market two (the market with constant returns to scale) was relatively elastic, then it is possible that the sign of $dK/da > 0$. Since the inelastic demand will keep the output for good one relatively constant, the derived demand for the joint input K will not fall very much. Market two will experience an increase in its derived demand for capital which

could offset any decrease from market one. Thus the overall change in the amount of capital could be positive.

Regulators should be aware of the fact that their decisions to change the cost allocation mechanism will have an effect on the size of the long run capital stock. If a regulator tries to increase the allocator to a division of a firm (say long distance) to benefit consumers of goods provided by another division of the firm (say local service) the regulatory action may cause the stock of the joint input (capital) to diminish. This could mean that the class of customers ostensibly benefitting from a reduced cost allocation may be hurt by the cost increases associated with a lower capital stock (that is, lower efficiency).

4.5 Conclusions

This chapter examined some simple cost allocation models which allocated the cost of a public input between the divisions of a fixed proportion production process firm. By examining even these simple models some of the output distortions anticipated by others (Averch-Johnson, 1961 and Needy, 1975) can be seen in a different light. Fixed proportions technology avoids the over-capitalization problem conjectured by Averch and Johnson because the firm can not alter its input mix to take advantage of a higher return on capital. By altering the allocation mechanism, there is an effect on the firms incentives to

chose the long run capital stock. In the long run with economies of scale in one division there was an incentive for the firm to reduce its capital stock as the allocating mechanism was increased.

The effect of a new allocator is not necessarily neutral, and may in fact cause total firm costs to rise in the long run. A reallocation of costs from (say) local service to (say) long distance could cause the size of the local network to diminish. This could have the effect of effectively increasing the cost of local service in terms of higher prices, or may impose costs in terms of lower efficiency on toll users.

4.6 End Notes

1. Without the assumption of positive production, one of the firm's equilibrium points is at zero production. Given the cost minimizing assumption, the firm could produce nothing and be minimizing costs. For this analysis it can be assumed that the firm must produce if it is in the market. The regulator can insure service by granting certificates of necessity to only those firms who are willing to serve positive units of output.

2. My thanks to Rich Romano for pointing this out and producing the graphical analysis.

CHAPTER V REGULATION AND COST ALLOCATION BETWEEN MARKETS

A cost allocation scheme's effect on the overall level of welfare for society depends upon the choices left to regulators and to the firm. For example, an unconstrained firm will choose a cost allocation mechanism that will allow profits to be maximized while a regulator might choose an allocation mechanism encompassing some idea of maximizing social welfare. The purpose of this paper is to examine incentives to the firm and industrial performance under various institutional arrangements and policy rules.

5.1 Cost Allocations between two Regulatory Jurisdictions

Hannon (1978) studied the cost allocation problem by examining the effects of the cost allocation mechanism on the output of the firm. The approach yields a great deal of information, but does not consider the problem from the input side. Hannon implicitly treated the shared capital (K) like a common input instead of a joint or public input in his model, with the firm subject to a regulatory constraint in both markets. Although, the model is similar to the standard Averch-Johnson (1962) models, Hannon points out the necessity to use these models,

If a minimally adequate theoretical base is to underlie the discussion of rate of return regulation, Averch-Johnson models must be expanded to encompass partial and multi-jurisdictional regulation ... the primary emphasis being on the impact of changes in the cost allocations between markets rather than analyzing or detecting A-J effects (Hannon, 1978: 31).

This point is especially valid given the well-known criticisms of Averch-Johnson models (Peles and Stein (1976), and Joskow (1974)).

Secondly, when Hannon examined the effects of a changing cost allocator (α) on the firm's behavior, he assumed that changing α had no effect on capital. As will be shown below, there could be a relationship between the cost allocation mechanism and the amount of capital used in the production process. There is no reason to assume a priori that Hannon's assumption is even true. In a more realistic setting (section 5.3), the effect of the cost allocation on capital becomes very important when examining the effect of a change in α on the choice of capital.

As mentioned previously, Hannon's model did not deal with joint costs, but with common costs (which can be defined loosely as those cost which are common to all production process). More precisely, however, when plant is an input in the production of two products, the cost of the input is a common cost if, in the production of one good, the amount of capacity left for production of other goods is diminished. There is, in fact, an opportunity cost of producing product one rather than product two.

Common costs could thus be allocated on the basis of cost causation since there is an opportunity cost of using the input for one particular product over another. With joint costs, however, there is no opportunity cost of producing one good over the other. In this sense too, a joint input is like a public input. In fact, while producing one output, the firm has available an input that can be used to produce another output.

The problem faced by regulators is that it is difficult to identify common costs and treat them differently than joint costs. The solution is often to lump the costs together and speak of allocating joint and common costs. In this respect Hannon's model simulates reality. Nevertheless, the firm still bases its decisions upon the marginal productivities of inputs even though it is subject to these accounting costs allocation rules.

Sweeny (1982) employed a model similar to the model described below using the accounting formulae examined in Braeutigam (1980). Sweeny examined the effect of two different fully distributed cost allocation formulas (relative output and relative revenues) on the output decisions of the firm given that it is subject to regulation in numerous jurisdictions. Sweeny's results showed that a firm using one of these accounting rules will choose a price vector which is dominated (i.e. there is a different set of price-output combinations that will yield

greater than or equal outputs with at least the same level of overall profit.) The model presented in the next sections examines the problem of allocating joint costs, and the effect of changing the cost allocation formula on output and capital input use. Unlike both Sweeny and Hannon, however, this model assumes that the input choice is effected by the cost allocation formula.

5.2. A Two Jurisdiction Model with Joint Costs

This section seeks to analyze the effects of altering the cost allocation factor (α) on the firm's input choice and output mix in the short run and in the long run. The comparative statics of a stylized model of cost allocation with regulation will be analyzed to show how much more complex the firm's decisions have become and to show how the firm, if left to maximize profit, will choose a cost allocation scheme consistent with the regulatory constraints.

The production function used in these models is a general neoclassical production function. The major difference here is that one input (K) enters the production function of two different goods. The models described below have a firm producing different outputs in two markets. Thus the production functions take the following form:

$$q_i(K, z_i) \text{ for } i = 1, 2,$$

where $q_K^i > 0$ and $q_z^i > 0$. The input K is a joint input into the production of both goods 1 and two while the input z_i can be a vector of other inputs unique to the production of good i .

The analysis next proceeds in four stages. First, the short-run, and thus, the variable cost functions are derived for all levels of the public input K . Second, profit will be maximized over the choice of K , q_1 , and q_2 subject to the regulatory constraint that price in each market be no greater than average cost. Third, we let the firm choose α , and fourth we will explore the implications of deviating from the firm's preferred α .

Now let $C^1(q_1) = V^1(q_1, K, w) + \alpha rK$, and

$$C^2(q_2) = V^2(q_2, K, w) + (1-\alpha)rK,$$

where $V^i(q_i, K, w) = \min_{z_i} w z_i$ s.t. $q_i^* = q_i(K, z_i)$ for $i = 1, 2$,

Forming the Lagrangean and deriving the first order conditions,

$$L z_i = w - \mu_i q_z^i = 0 \quad i = 1, 2, \quad (5.2.1)$$

$$L \mu_i = q_i^* - q_i(K, z_i) = 0 \quad i = 1, 2, \quad (5.2.2)$$

where μ_i is the Lagrangean multiplier for market i .

By employing the implicit function theorem, the conditional demand functions for each of the variable inputs can be derived. Thus, $z_i^* = z_i^*(q_i, K, w)$ for each output and the variable cost functions V^i are obtained so that

$$v^i(q_i, K, w) = wz_i^* \quad i = 1, 2. \quad (5.2.3)$$

This variable cost function thus has the following properties:

$$i) \quad v^i_1 = (\delta v^i / \delta q_1) > 0, \text{ for } i = 1, 2.$$

$$ii) \quad v^i_K = (\delta v^i / \delta K) < 0, \text{ for } i = 1, 2.$$

$$iii) \quad v^i_K = w^*(\delta z_i / \delta K)$$

The first property is that marginal (variable) costs must be positive for both outputs. The second property states that variable costs decrease as the amount of capital increases. The third property, derived from the definition of the variable cost function, states that the derivative of the variable cost function with respect to K is the marginal rate of technical substitution between z_i and K multiplied by the wage rate (w) of z_i .

As a bench-mark to compare with the next model let us assume we have the following total cost function:

$$C(q_1, q_2) = v^1(q_1, K) + v^2(q_2, K) + rK \quad (5.2.4)$$

Given this function, we now desire to determine the optimal provision conditions for K . From the public goods literature (Samuelson, 1956) we know the firm will choose K to minimize costs by setting the summation of the marginal rates of technical substitution equal to the ratio of the wage rates. This can be done by taking the derivative of the above cost function with respect to K so as to yield the following first order condition:

$$v_K^1 + v_K^2 = r \quad (5.2.5)$$

Rearrangement yields:

$$(\delta z_1 / \delta q_1) + (\delta z_2 / \delta q_2) = r/w \quad (5.2.6)$$

Now that the technology can be described, the next step is to maximize profit (π) over the choice of K , q_1 , and α , given the regulatory constraint that prices must be no greater than average cost as defined by the regulators. This constraint is the familiar rate of return constraint which allows the firm a specified rate of return on capital. If r is the true economic cost of capital, the regulators pick a $s_1 > r$ as a cap on the firm's earnings. Therefore, price in the first jurisdiction must be:

$$p_1 \leq (v^1 + \alpha s_1 K) / q_1,$$

and price in jurisdiction two must be less than or equal to average costs as defined by:

$$p_2 \leq (v^2 + (1 - \alpha) s_2 K) / q_2.$$

The firm is allowed to choose its level of output, capital input use, and cost allocation scheme. To do so, it undertakes the following maximization:

$$\begin{aligned} \text{Max } \pi &= R^1(q_1) + R^2(q_2) - v^1(q_1, K, w) - v^2(q_2, K, w) - rK \\ &\text{K, } q_1, \alpha \\ \text{s.t. a) } &R^1(q_1) - v^1(q_1, K, w) - \alpha s_1 K \leq 0 \\ \text{b) } &R^2(q_2) - v^2(q_2, K, w) - (1 - \alpha) s_2 K \leq 0 \\ \text{c) } &q_i > 0, K > 0, \text{ for } i = 1, 2. \end{aligned}$$

where K represents the long run demand for K .

Before analyzing the Kuhn-Tucker first order conditions, it is necessary to make some assumptions about the

revenue and cost functions. For a maximum, the Kuhn-Tucker sufficiency theorem requires the objective function to be differentiable and concave in the non-negative orthant. In addition, each constrained function must be differentiable and convex in the non-negative orthant. According to the Kuhn-Tucker sufficiency condition, if a solution satisfies the Kuhn-Tucker first order conditions it will be a global maximum (Chiang, 1974: 722). In addition, the following regularity conditions are assumed:

$$i) \quad \frac{\delta R^i}{\delta q_i} = R_i^i > 0 \quad i = 1, 2.$$

$$ii) \quad \frac{\delta^2 R^i}{\delta q_i \delta q_i} = R_{ii}^i < 0 \quad i = 1, 2.$$

$$iii) \quad \frac{\delta V^i}{\delta q_i} = V_i^i > 0 \quad i = 1, 2.$$

$$iv) \quad \frac{\delta^2 V^i}{\delta q_i \delta q_i} = V_{ii}^i < 0, \quad i = 1, 2.$$

$$v) \quad \frac{\delta V^i}{\delta K} = V_K^i < 0, \quad i = 1, 2.$$

$$vi) \quad \frac{\delta^2 V^i}{\delta K \delta K} = V_{KK}^i > 0, \quad i = 1, 2.$$

$$vii) \quad \frac{\delta^2 V^i}{\delta K \delta q_j} = 0, \quad i \neq j.$$

$$viii) \quad \frac{\delta^2 V^i}{\delta q_i \delta q_j} = 0, \quad i \neq j.$$

The first four conditions are the typical first and second order conditions. Conditions (v) and (vi) are the same as those derived above during the discussion of the variable cost function. The last two conditions though are used to get at the public good nature of the capital input. Condition (vii) says that a change in output in one market has no effect on the cost savings attributable to K in the other market. This "independence" assumption is employed to make the analysis a bit less complicated. Condition (viii) requires that the marginal cost of one market be independent of a change in output in the other market. This will preclude the possibility of congestion in the use of the joint input.

Letting $M(q_1, q_2, K, \alpha)$ be the objective function, the Kuhn-Tucker first order conditions are:

$$M_K = -V_K^1(1-\mu_1) - V_K^2(1-\mu_2) - r + \alpha s_1 \mu_1 + (1-\alpha)s_2 \mu_2 \leq 0 \quad (5.2.7a)$$

$$M_{KK} = 0 \quad (5.2.7b)$$

$$M_1 = (R_1^1 - V_1^1)(1 - \mu_1) \leq 0 \quad (5.2.8a)$$

$$M_1 q_1 = 0 \quad (5.2.8b)$$

$$M_2 = (R_2^2 - V_2^2)(1 - \mu_2) \leq 0 \quad (5.2.9a)$$

$$M_2 q_2 = 0 \quad (5.2.9b)$$

$$M_\alpha = (s_1 \mu_1 - s_2 \mu_2)K \leq 0 \quad (5.2.10a)$$

$$M_\alpha \alpha = 0 \quad (5.2.10b)$$

$$M_{\mu_1} = -R^1 + V^1 + \alpha s_1 K \geq 0 \quad (5.2.11a)$$

$$M_{\mu_1 \mu_1} = 0 \quad (5.2.11b)$$

$$M_{\mu_2} = -R^2 + V^2 + (1-\alpha)s_2 K \geq 0 \quad (5.2.12a)$$

$$M_{\mu_2\mu_2} = 0 \quad (5.2.12b)$$

Given these first order conditions it is necessary to rule out values the choice variables may take in order to be assured a solution exists.

From the examination of the constraints, two feasible solutions exists. The values of the choice variables which are consistent with the first solution are:

$$K > 0, \mu_1 = 0, \mu_2 > 0, q_i > 0, \text{ for } i = 1, 2.$$

This is the case where the firm is regulated (constrained) in only one market (market 2). From the constraint analysis the only possible value of α which can exist without a violation of any of the constraints is $\alpha = 0$. This implies that none of the joint costs are allocated to the unregulated market. If α were greater than 0, then the constraint in equation (5.2.10a) should be zero by complementary slackness. But since $\mu_1 = 0$, equation (5.2.10a) implies $-\mu_2 s_2 = 0$. However, since it was assumed μ_2 and s_2 are positive, the constraint is violated. Therefore α must be zero in this solution.

Temin and Peters (1985), in their explanation of the history of the separations process, conclude that it was in AT&T's best interests to put all of its joint plant in the local jurisdiction before the FCC took an active stance in the regulation of interstate messages. The above model partly explains why AT&T fought the introduction of the station to station method of accounting where AT&T was

required to charge a portion of its joint plant used in the production of toll and local calls to both the long distance and local jurisdictions. This Kuhn-Tucker solution is essentially the solution with the board to board method of accounting: the allocation scheme favored by AT&T prior to FCC regulation. The Kuhn-Tucker multiplier μ_1 equal to zero implies that there was no effective regulatory constraint, nor was there any reason to share costs if they could be allocated completely to the regulated jurisdiction.

Similarly, when the FCC attempted to split AT&T into two separate subsidiary companies in the Computer II (1980) decision (one regulated and the other unregulated) to allow the unregulated long distance division to compete against unregulated firms on a "fair" basis, the FCC recognized the incentive to shift its common and joint costs to the regulated local jurisdiction. In order to avoid this result the FCC instituted very detailed and complicated rules to make sure the integrated telecommunications firm would not be expanding its markets at the expense of the local telephone customer or have an unfair advantage over its competitors.

Returning to the model, we note that the second solution is the case where the firm is subjected to regulation in both markets ($\mu_1, \mu_2, > 0$), a positive amount of capital is chosen ($K > 0$), and output in both markets is

positive ($q_1 > 0$). The cost allocation mechanism (α) consistent with all the first order conditions, not surprisingly, can range from zero to one depending upon the system of the first order conditions.

Examining the first order condition in equation (5.2.7a) while assuming that $K > 0$, let us compare this result with the benchmark cost minimizing condition shown in equation (5.1.5). By rearranging (5.2.7a) and defining

$v_K^i (1 - \mu_i) \equiv -wz_K^i$, for $i = 1, 2$, then

$$\hat{z}_K^1 + \hat{z}_K^2 = (r - \alpha s_1 \mu_1 - (1 - \alpha) s_2 \mu_2) / w \quad (5.2.13)$$

Equation (5.2.13) says that the optimal amount of K is chosen based upon the constrained sum of the marginal rates of technical substitution (MRTS) between K and z_i being set equal to a function of the ratio of the wage rates. From the cost minimizing case in equation (5.2.2) we can see that the choice of K was based simply on the sum of the unconstrained MRTS being set equal to the ratio of the wage rates.

Now, if α is chosen by the firm and is positive, then, from equations (5.2.10a) and (5.2.10b), we know that

$$(\mu_1 s_1 - \mu_2 s_2) = 0.$$

Thus, (5.2.13) reduces to

$$\hat{z}_K^1 + \hat{z}_K^2 = (r - s_1 \mu_1) / w \quad (5.2.13a)$$

Not surprisingly, the expression on the right hand side of (5.2.13a) is not equal to the ratio of wage rates (r/w) and in fact is greater than (r/w) since $\mu_1 s_1 > 0$. This is the

standard Averch-Johnson result. Only in this case the input is a joint input into the production of both services. By allowing the firm to choose α , the firm will pick an α that will equate the effective regulatory constraints across the jurisdictions thus yielding the standard AJ result.

After examining the constraint and finding a solution, it is possible to derive the values of the choice variables that maximize profits. Using the implicit function theorem the optimal values are:

$$q_i^*(w, r, s_i) \quad i = 1, 2.$$

$$K^*(w, r, s_i)$$

$$\alpha^*(w, r, s_i)$$

Thus, given a rate of return constraint requiring prices to be no greater than allowed average cost, and allowing the firm to choose its capital input and output in both markets, the firm will choose an α to allocate costs between the two divisions in such a way that is consistent with the regulatory constraints.

It is important to note that if regulators in both jurisdictions set $s_i = r$, the firm will act to set prices consistent with the new allowed rates of return. What is interesting about these prices is that they will be Ramsey prices. Equations (5.2.8a) and (5.2.9a) can be rewritten so that:

$$\frac{(p_1 - v_1^1)}{p_1} e_1 = \frac{(p_2 - v_2^2)}{p_2} e_2 \quad (5.2.14)$$

From equations (5.2.11a) and (5.2.12a) we know that price must equal average cost which is a function of α . By inserting these prices (p_i^*) into (5.2.14) we get:

$$\frac{(p_1^* - v_1^1)}{p_1^*} e_1 = \frac{(p_2^* - v_2^2)}{p_2^*} e_2 \quad (5.2.15)$$

These prices are equal to the 'average cost' when α is chosen to set $\mu_1 = \mu_2$ from the first order condition in equation (5.2.10a). This has the effect of constraining the firm to just one regulatory constraint which results in the standard Ramsey result. Even when $s_1 \neq r$, a "Ramsey like" result will occur because the firm attempts to set α so as to equate $\mu_1 s_1$ and $\mu_2 s_2$. The result, which is not even a second best result, will depend upon the relative demand elasticities.

Theoretically, it is not necessary for regulators to set α since regulators implicitly determine α when they set the allowed rates of return. If $s_1 > s_2$ we would expect the firm to pick a particular α , and if the sign was reversed, the firm would choose another allocation. However, the reality of the situation is that the cost allocation factor, α , is a politically determined variable. The history of the separations process detailed in Gabel

(1967), Sichter (1977), and Temin and Peters (1985) demonstrates that the telecommunications cost allocator had changed numerous times in the last four decades to rectify imbalances in interstate versus intrastate toll rates, to stave off rate increases for the local companies, and to keep AT&T from earning excess profits. The next logical question to ask is what happens to the output choices and level of capital chosen if the allocating factor is changed by the regulators to meet their policy goals.

5.3 Long Run Results when α is Exogenous

As mentioned above, Hannon (1978) and Sweeney (1982) either assumed or implicitly set $dK/d\alpha = 0$. In so doing, they were unable to examine the effect of exogenous changes in α on all the firm's decisions. By taking the total differential of the Kuhn-Tucker first order conditions (assuming an exogenous α set by regulators) a linear system of five equations is derived and shown in Table 5.1. Now, by dividing both sides of the system by $d\alpha$ and rearranging the system into matrix notation such that the system is now represented as $Hx = d$ where

$$H = \begin{vmatrix} 0 & 0 & M_{u1K} & M_{u1q1} & 0 \\ 0 & 0 & M_{u2K} & 0 & M_{u2q2} \\ M_{u1K} & M_{u2K} & M_{KK} & M_{K1} & M_{K2} \\ M_{u1q1} & 0 & M_{K1} & M_{11} & 0 \\ 0 & M_{u2q2} & M_{K2} & 0 & M_{22} \end{vmatrix}$$

$$\text{and where } x = \begin{vmatrix} du_1/d\alpha \\ du_2/d\alpha \\ dK/d\alpha \\ dq_1/d\alpha \\ dq_2/d\alpha \end{vmatrix}, \text{ and } d = \begin{vmatrix} -s_1K \\ s_2K \\ 0 \\ 0 \\ 0 \end{vmatrix}.$$

Using Cramer's rule to solve for the elements of x , the comparative statics can be derived and are shown in Table 5.2.

As can be seen in Table 5.2, the signs of the total derivatives with respect to α are not absolutely determinable, but it is still possible to obtain some information. The fact that $dK/d\alpha$ is unlikely to be zero is important in and of itself. By examining the denominator (D) we know that for a maximum, the sign of the principal minors must alternate and the sign for the determinant of the bordered Hessian must be negative. Although not shown here, this condition has been met. Equation (1) in Table 5.2 shows the comparative static for $dK/d\alpha$. The first term ($B_1M_{22} - B_2M_{K2}$) < 0 since $M_{22} < 0$, $M_{K2} > 0$ and $B_1 > 0$. Correspondingly, the second term is positive since it is preceded by a negative sign.

Before going further, the importance of differing regulatory policies should be noted in conjunction with the AJ effect. There is still the mathematical result that a firm will choose to employ more than the optimal amount of capital when allowed to earn a rate of return greater than the cost of capital. Even if some of the behavioral implications of the AJ hypothesis are subject to debate, there is the added problem of how the capital choice is affected when the firm employs capital subject to cost allocation regulation. This new distortion depends in part

Table 5.1
Total Differentiation of First Order Conditions

$$\begin{aligned}
 & + \pi_{\mu 1, K} dK + \pi_{\mu 1, q1} dq_1 = -s_1 K da \\
 & + \pi_{\mu 2, K} dK + \pi_{\mu 2, q2} dq_2 = s_2 K da \\
 \pi_{\mu 1, K} d\mu_1 + \pi_{\mu 2, K} d\mu_2 + \pi_{KK} dK + \pi_{K1} dq_1 + \pi_{K2} dq_2 &= (s_2 \mu_2 - s_1 \mu_1) da \\
 \pi_{\mu 1, q1} d\mu_1 + \pi_{K1} dK + \pi_{11} dq_1 &= -s_1 \mu_1 K_1 da \\
 \pi_{\mu 2, q2} d\mu_2 + \pi_{K2} dK + \pi_{22} dq_2 &= s_2 \mu_2 K_2 da
 \end{aligned}$$

$$\begin{aligned}
 \pi_{KK} &= -\Sigma V_{KK}^1 (1 - \mu_1) < 0 & \pi_{\mu 1, K} &= V_K^1 + as_1 > 0 \\
 \pi_{K1} &= -V_{K1}^1 (1 - \mu_1) > 0 & \pi_{\mu 1, q1} &= -(R_1^1 - V_1^1 - as_1 K_1) < 0 \\
 \pi_{K2} &= -V_{K2}^2 (1 - \mu_2) > 0 & \pi_{\mu 2, K} &= V_K^2 + (1 - \alpha)s_2 > 0 \\
 \pi_{\mu 2, q2} &= -(R_2^2 - V_2^2 - (1 - \alpha)s_2 K_2) > 0 \\
 \pi_{11} &= [(R_{11}^1 - V_{11}^1)(1 - \mu_1) + K_{11}(as_1 \mu_1 - r)] < 0 \\
 \pi_{22} &= [(R_{22}^2 - V_{22}^2)(1 - \mu_2) + K_{22}((1 - \alpha)s_2 \mu_2 - r)] < 0
 \end{aligned}$$

$$\begin{aligned}
 (1) \quad \frac{dK}{d\alpha} &= \frac{1}{D} \left(K(-\lambda_2 B_2(\mu_1 s_1 - \mu_2 s_2)) \right) - \underbrace{s_2 \lambda_2 (B_1 M_{22} - B_2 M_{K2})}_{?} - s_1 B_2 (\lambda_1 M_{11} - \lambda_2 M_{K1}) \stackrel{+}{\leq} 0 \\
 (2) \quad \frac{dq_1}{d\alpha} &= \frac{1}{D} \left(K(\lambda_1 \lambda_2 B_2^2 (\mu_1 s_1 - \mu_2 s_2)) \right) + \underbrace{s_2 \lambda_1 \lambda_2 (M_{K3} B_2 - M_{22} B_1)}_{?} + s_1 (\lambda_1 B_1^2 M_{K1} + 2 \lambda_2 B_1 B_2 M_{K2} - \lambda_2 B_1^2 M_{22} - \lambda_2 B_2^2 M_{KK}) \stackrel{+}{\geq} 0 \\
 (3) \quad \frac{dq_2}{d\alpha} &= \frac{1}{D} \left(K(\lambda_2^2 B_1 B_2 (\mu_1 s_1 - \mu_2 s_2)) \right) - \underbrace{s_1 B_1 B_2 (M_{K1} \lambda_2 - M_{11} \lambda_1)}_{?} - s_2 (\lambda_2^2 B_1 M_{K2} + 2 \lambda_1 \lambda_2 B_2 M_{K1} - \lambda_2^2 B_2 M_{KK} - \lambda_2^2 B_2 M_{11}) \stackrel{-}{\geq} 0
 \end{aligned}$$

$$(4) \quad D = B_1^2 \lambda_2^2 M_{22} - 2 B_1 \lambda_2 B_2^2 M_{K2} - 2 M_{K1} \lambda_1 \lambda_2 B_2^2 + \lambda_2^2 B_2^2 M_{KK} + M_{11} \lambda_1^2 B_1^2 < 0$$

$$B_1 = \frac{\delta^2 M}{\delta K \delta \mu_2} > 0 \quad A_1 = \frac{\delta^2 M}{\delta K \delta \mu_1} > 0$$

$$B_2 = \frac{\delta^2 M}{\delta q_2 \delta \mu_2} > 0 \quad A_2 = \frac{\delta^2 M}{\delta q_2 \delta \mu_1} > 0$$

Table 5.2
Comparative Statics

upon the relative size of $\mu_1 s_1$ versus $\mu_2 s_2$, or divergent regulatory treatment. This behavior may, in fact, be a more plausible reason for any capital use distortions or output distortions than the standard AJ distortion.

Now looking at equation (2) in Table 5.2 the sign of dq_1/da depends upon the relative importance of the regulatory constraints as does the sign of dK/da . The sign of the second and third terms are able to be determined and turn out to be positive; so, if $(\mu_1 s_1 - \mu_2 s_2)$ is positive (or relatively small) then the sign of dq_1/da in equation (2) will be negative since the numerator is positive and the denominator is negative. Similarly, in equation (3) the second and third terms are negative. Coupled with the possibility that $(\mu_1 s_1 - \mu_2 s_2)$ is negative (or positive and relatively small), the sign of dq_2/da in equation (2) will be positive.

For $dq_1/da > 0$ in equation (2), and $dq_2/da < 0$ in equation (3) the effect of the differences between the effective regulatory constraints $(\mu_1 s_1 - \mu_2 s_2)$ must outweigh the effects of the terms that have a definite sign. The fact that the sign of the entire differential equation can not be determined with certainty suggests that the possibility that a perverse policy result could occur. For example, say regulators increased α (thus increasing the share of the joint-costs to market one) to encourage more consumption of good two by making the price of good

two lower. Then if, $dq_2/da < 0$ because of the relative effects of the effective regulatory constraint ($\mu_1 s_1 - \mu_2 s_2$) terms outweighing the terms with a definite sign, the increase in a will have the exact opposite effect from what the policy intended. Because of this possibility, however remote it might be, regulators should take care to ascertain the effects of changing the allocator.

Finally, the sign of da/dK will depend, to a great extent, on the relative demand elasticities for the two markets. If demand in market 1 is relatively inelastic, the MRTS will be insensitive to changes in price, and if market 2's demand is elastic, we would expect to see an increase (shift) in the MRTS for market 2 which would offset any decrease in market 1's MRTS.

5.4 Short Run Analysis with an Exogenous Change in a

In the short run, when the firm cannot alter its underlying capital structure, it is interesting to note the effects of an exogenous change in the cost allocation formula on the output in the two markets. Unlike the long run results, more definitive results can be derived. As mentioned above, the cost allocation mechanism has been altered in the past for numerous policy reasons. If the regulator were to arbitrarily increase the cost allocation mechanism, the firm will shift some of the costs presently covered by one division to the other division of the firm. This would cause prices to change in order to be consistent

with the regulatory constraints. By setting $dK = 0$ it is possible to see the short run affects of a change in α .

Solving the system $H_1 x_1 = x_1$ for x_1 where:

$$H_1 = \begin{vmatrix} 0 & 0 & -(R_1^1 - V_1^1) & 0 \\ 0 & 0 & 0 & -(R_2^2 - V_2^2) \\ -(R_1^1 - V_1^1) & 0 & M_{11} & 0 \\ 0 & -(R_2^2 - V_2^2) & 0 & M_{22} \end{vmatrix}$$

$$x_1 = \begin{vmatrix} d\mu_1/d\alpha \\ d\mu_2/d\alpha \\ dq_1/d\alpha \\ dq_2/d\alpha \end{vmatrix}, \text{ and } d_1 = \begin{vmatrix} -s_1K \\ s_2K \\ 0 \\ 0 \end{vmatrix},$$

the following comparative statics can be derived.

$$\frac{d\mu_1}{d\alpha} = \frac{s_1(KM_{11})}{(R_1^1 - V_1^1)^2} < 0 \quad (5.4.1)$$

$$\frac{d\mu_2}{d\alpha} = \frac{s_2(-KM_{22})}{(R_2^2 - V_2^2)^2} > 0 \quad (5.4.2)$$

$$\frac{dq_1}{d\alpha} = \frac{s_1K}{(R_1^1 - V_1^1)} < 0 \quad (5.4.3)$$

$$\frac{dq_2}{d\alpha} = \frac{-s_2K}{(R_2^2 - V_2^2)} > 0 \quad (5.4.4)$$

Equation (5.4.1) is negative since $M_{11} < 0$. Similarly, for equation (5.4.2), $M_{22} < 0$, so $d\mu_2/d\alpha > 0$. These two equations are difficult to interpret, but since the Kuhn-Tucker multiplier is essentially the marginal profitability of the division, equation (5.4.1) says that if α were increased, the marginal profitability of the firm will decrease. For division two, equation (5.4.2) says that as α increases the marginal profitability of the firm will increase. This makes intuitive sense because the division's costs are being administratively altered, so

there would be an effect upon the division's marginal profitability.

For the output markets, as α increases the amount of q_1 supplied by the firm decreases reflecting an increase in costs. To satisfy the regulatory constraint, the new allocator requires an increase in prices and a decrease in quantity demanded. Equation (5.4.3) shows this result assuming that a regulated firm produces past the point where marginal revenue equals marginal costs (that is, $R_1^i < V_1^i$). In market two, described in equation (5.4.4), as α increases costs fall to the division, and the firm is required to increase its production of q_2 to satisfy the profit constraint.

In the short run when the capital input is fixed, the regulator can influence the relative amount of output produced. The numerous changes made by the regulatory authorities from 1947 to 1982 can plausibly be explained by the simple fact that only short run considerations were taken into account by policy makers. The fact that a firm would adjust its production decisions (potentially in opposition to the regulator's policies) in the long run was never considered. If regulators did not understand the concepts embodied in a demand elasticity until the early 1960s, then it seems highly plausible that a supply response based upon these altered cost allocations would not be considered either. Although, it is not possible to

describe the long run effect on the firm without an idea of the firm's capital responsiveness to a change in the allocation, it is important to know that such a response is likely. As will be shown in the next section, this responsiveness can provide some interesting results in a welfare analysis context.

5.5 An Application: Welfare and Ramsey Prices

This section shows how the cost allocation chosen by the regulator can be compared to the Ramsey optimal choice. It is possible to test the hypothesis that regulators have preferred one class of customers over another by examining how regulators chose the cost allocation (α). But, first, a short digression is in order to show how a similar technique was used to make conclusions concerning the relative social welfare weights a regulator may place on the value of consumption of different goods of a multiproduct monopolist.

5.5.1 Implied Social Welfare Weights

Ross (1984) has shown that it is possible to retrieve an implied social welfare weighting made by regulators between different goods. Ross asks whether it is possible to infer from existing price structures the social welfare weights that would make prices Ramsey-Optimal. To see this, assume there are two classes of customers where A customers consume only good 1 and B customers consume only good 2, and that there is an over all profit

constraint on the firm. This gives the following maximization problem for the firm:

$$\max_{p_i} W = \Gamma^A Z^A(p) + \Gamma^B Z^B(p) + \Theta(\pi - \pi^*) \quad (5.5.1)$$

and gives the following first order conditions:

$$W_1 = -q_1^A \Gamma^A + \Theta(q_1 + (p_1 - c_1) \frac{dq_1}{dp_1}) = 0 \quad (5.5.2)$$

$$W_2 = -q_2^B \Gamma^B + \Theta(q_2 + (p_2 - c_2) \frac{dq_2}{dp_2}) = 0 \quad (5.5.3)$$

where $Z^A(p)$ and $Z^B(p)$ are the consumers' surplus for consumers A and B, Γ^A and Γ^B are the regulator's weights for each class of consumers, and Θ is the value the firm's profits takes in the regulator's social welfare function. From Shephard's Lemma we also know that $\delta Z / \delta p_i = -q_i$.

Dividing the FOC's by q_i , and using a slight manipulation we get:

$$\frac{(p_1 - c_1)}{p_1} e_1 = \frac{\Theta - \Gamma^A}{\Theta} \quad (5.5.4)$$

$$\frac{(p_2 - c_2)}{p_2} e_2 = \frac{\Theta - \Gamma^B}{\Theta} \quad (5.5.5)$$

These are the Ramsey results when there are differing social welfare weights. If the weight for A equaled the weight for B, then we would have the traditional formulation where the percentage price mark-up over marginal costs

multiplied by the price elasticity is a constant across all markets.

Now by rewriting (5.5.4) and (5.5.5) as

$$r_A/\theta = 1 - \frac{(p_1 - c_1)}{p_1} e_1 = 1 - m_1 e_1, \quad (5.5.4a)$$

$$r_B/\theta = 1 - \frac{(p_2 - c_2)}{p_2} e_2 = 1 - m_2 e_2, \quad (5.5.5a)$$

it is possible to derive a simple expression relating the welfare weights by dividing the first equation by the second equation. This yields:

$$\frac{r^A}{r^B} = \frac{1 - m_1 e_1}{1 - m_2 e_2}. \quad (5.5.6)$$

To find the relative weights we need only the mark-up over price and the price elasticity of demand for each good. Then it is possible to econometrically test to see if r^A/r^B is significantly different than one to determine whether the class of A customers are favored over the class of B customers.

5.5.2 Determining the Optimal α

Now let us introduce a slightly different formulation which is consistent with the model described in section 5.4. If we assume that there are two regulatory jurisdictions and that there is a required cost allocation between the two jurisdictions, then the maximization problem becomes:

$$\begin{aligned} \max_{p, \alpha} W = & Z^A(p_1) + Z^B(p_2) + \mu_1[\pi_1^* - R^1 - V^1 - \alpha rK] + \\ & + \mu_2[\pi_2^* - R^2 - V^2 - (1-\alpha)rK] \quad (5.5.7) \end{aligned}$$

This problem yields the following first order conditions.

$$W_1 = \delta Z^A / \delta p_1 - \mu_1(q_1 + p_1 \delta q_1 / \delta p_1 - V_1^1 \delta q_1 / p_1) = 0 \quad (5.5.8)$$

$$W_2 = \delta Z^B / \delta p_2 - \mu_2(q_2 + p_2 \delta q_2 / \delta p_2 - V_2^2 \delta q_2 / p_2) = 0 \quad (5.5.9)$$

$$\begin{aligned} W_\alpha = & \Sigma \delta p_i / \delta \alpha [\delta Z^j / \delta p_i - \mu_i(q_i + p_i \delta q_i / \delta p_i - V_i^i \delta q_i / p_i)] \\ & + rK(\mu_1 - \mu_2) = 0. \quad (5.5.10) \end{aligned}$$

where $j = A, B$, and $i = 1, 2$. Since equations (5.5.8) and (5.5.9) are equal to 0, equation (5.5.10) reduces to $\mu_1 = \mu_2$. Thus by choosing the optimal level of α the regulators equate the Lagrangean multipliers across jurisdictions. This is the same result we would get if the firm chose the cost allocator. Thus, by choosing α optimally, the regulator effectively sets Ramsey prices. To the extent that α is not chosen to satisfy (5.5.10) one could econometrically test to see how far regulatory policies diverge from the Ramsey efficient result. By rearranging equations (5.5.8) and (5.5.9) using the steps outlined in equations (5.5.4) - (5.5.6) we can see that

$$\frac{\mu_1^{-1}}{\mu_1} = (p_1 - V_1)e_1/p_1 = (p_2 - V_2)e_2/p_2 = \frac{\mu_1^{-1}}{\mu_1}.$$

This in turn implies that

$$1 = \frac{1 - m_1 e_1}{1 - m_2 e_2} . \quad (5.5.11)$$

It is thus possible to test whether equation (5.5.11) is greater than, less than, or equal to one. Notice that this is the same test that Ross proposed in equation (5.5.6) to determine whether regulators were employing social welfare weights when pricing decisions. Thus from the same evidence Ross would use to extract social welfare weights it is possible to make conclusions, not about welfare weights, but about the optimal cost allocation. To the extent that (5.5.11) is (say) greater than 1 it would be difficult to say with confidence that regulators chose to weight the welfare of one group of consumers over another since an alternative conclusion would be that the regulators chose a sub-optimal cost allocator. Given the fact that regulators do not always attempt to set Ramsey prices either due to lack of knowledge or a lack of technical ability, then the second conclusion that an incorrect cost allocation has been chosen is more applicable.

5.6 Welfare and Passive Regulation

As mentioned above, the FCC and state public service commissions have used the cost allocation mechanisms to transfer rents between local and toll jurisdictions. This regulatory behavior is not the traditional rate base

regulation found in the Averch-Johnson line of models. In fact, the behavior is more like that modeled by Joskow (1974). Joskow hypothesized that regulatory agencies seek to maximize their utility by minimizing conflict and criticism appearing in the regulatory environment. Consumer groups, legislators, investors, and the regulated industries themselves generate signals to the regulators to which the regulators respond. Regulation, under Joskow's approach, is characterized by an "equilibrium relationship" where regulatory procedures and instruments are well defined and are used repetitively and predictably.

Joskow hypothesized that during the 1950's and 1960's the regulation of the electric utility industry was in this equilibrium mode as regulators were most concerned with keeping nominal prices from increasing. Electricity prices declined as costs declined. Firms which could increase their profitability without increasing prices or allowing overall profits to go above a "reasonable" level were not subjected to strict rate base regulation. Rate decreases were regularly allowed without investigation, and it was only when a firm asked for a rate increase did the regulators inquire into the firm's costs and revenues. Thus, the regulation was passive and firms were permitted to earn almost any rate of return as long as it did not go above an "unreasonable" amount (see *FPC v. Hope Natural Gas*,

1944). So, there was no "allowed" rate of return which regulators forced the companies to earn.

Although Joskow only used examples of the electric industry to illustrate his theory, the telecommunications industry also faced similar passive regulation. Local companies were not subject to strict scrutiny unless asking for rate increases. Regulators never forced the local companies to give back excess earnings until very recently. Rate reviews were only called for when the local company asked to raise its rates. It was only when enough states had companies asking for rate increases was the cost allocation between local and toll calls altered so that the local company rates no longer needed an increase. Costs were shifted from the intrastate jurisdiction to the interstate jurisdiction through the allocation change to avoid increases in nominal prices.

This model is consistent with the joint state-federal regulation of the telephone industry. It should be noted that this joint regulation was not really an amicable arrangement. The states and federal government each had different constituencies. Although not as passive as the states, the FCC negotiated with AT&T to reduce rates from time to time. The state governments resented these rate reductions because it made the rates for interstate toll calls cheaper than those placed within-state. This had the effect of increasing the criticism the state

regulators received. The state utility regulators banded together to fight the FCC and wrestle away some of the FCC's jurisdiction. In order to avoid losing the fight, the FCC (to some extent) allowed the states (in the form of the National Association of Regulated Utility Commissioners) some say over how costs are allocated (Temin and Peters 1985). Table 5.3 shows the cost allocation to the interstate jurisdiction over time. Since 1947, the cost allocation of the joint plant to the interstate jurisdiction has increased from approximately 5 percent to 25 percent in 1984. Gabel (1967) cites numerous examples throughout his monograph of changes in the cost allocation supposedly to stave off local rate increases.

If the Joskow approach is valid and the cost allocation was altered to keep local rates from increasing, the explanation as to why regulators were able to shift these costs with apparent impunity were twofold. First, costs were falling in the long distance market due to a combination of economies of scale and technological change. Regulators could "tap" the rents from the long distance market (without raising nominal prices in the long distance market) and apply them to the local jurisdiction to keep rates from increasing. Second, there were no long distance telecommunications substitutes, so long distance consumers were not able to change telecommunications suppliers.

If we assume the firm maximizes profit subject to some type of profit constraint like average cost plus a mark-up. It could even be the traditional AJ type of regulation described in the models in section 5.2. The only necessary assumption is that the regulator only acts to change the cost allocation formula when one of the firm's divisions meets a binding regulatory constraint and thus comes before the commission to ask for a change in the allowed rate of return. If the constraint becomes binding the firm must get approval from the regulatory commission to increase its rates. But, instead of granting a increase in the division's mark-up, the regulators shift some of the burden of the joint cost to the other jurisdiction.

The regulators problem can be examined under two sets of regulatory behavior. The first is when regulators concern themselves only with the effects of their policies on the consumer. The second way of looking at this problem is when the regulators, in addition to being concerned about the consumers surplus, are also concerned about the firms profits.

For the first approach, assume that there are two classes of customers toll (T) and local (L). Local customers only purchase good 2 which is local service and toll customers only purchase good 1 which is toll service. The regulator is thus concerned solely with the effect of changing the cost allocation scheme on the sum of the

consumer's surplus of these two classes of customers (i.e. the regulator perceives the welfare weight on the firms profits to be zero). Let B^L and B^T be the benefit functions for customers of local and toll services respectively. Given that the regulators are going to change α , the maximization takes the following form:

$$\max_{\alpha} W(q_1, q_2, \alpha) = B^T(q_1(\alpha)) + B^L(q_2(\alpha)) \quad (5.6.1)$$

where W is the social welfare function regulators are trying to maximize. The first order condition is:

$$w_{\alpha} = \frac{\delta B^T}{\delta q_1} \frac{\delta q_1}{\delta \alpha} + \frac{\delta B^L}{\delta q_2} \frac{\delta q_2}{\delta \alpha} = 0 \quad (5.6.2)$$

This can be rearranged so that:

$$\frac{\delta B^T}{\delta q_1} \frac{\delta q_1}{\delta \alpha} = - \frac{\delta B^L}{\delta q_2} \frac{\delta q_2}{\delta \alpha} = 0 \quad (5.6.3)$$

Equation (5.6.3) says that regulators alter α so that the marginal benefits from consumption of these two goods is equated. As α increases, consumer surplus falls in the toll market and increases in the local market. What this type of regulation does is take some of the surplus from toll consumers and transfer it to local consumers so as to equate the social marginal benefit across markets.

A more realistic view of the regulatory process could be described as follows. Through some objective determination regulators attempt to set the allowed rate of return so that is comparable to similarly situated assets. The firm is thus allowed to set "reasonable"

prices for each output and chose its level of joint input. The regulators do not affirmatively act to set rates, and only investigates the firm's revenue and costs when the firm asks for a rate increase. The regulator's response in this model when confronted with a request for a rate increase is to alter the cost allocation between markets rather than allow prices to increase in the local jurisdiction. The regulator's choice variable is thus the cost allocator.

Now by adding the firm's maximized profit function to the social welfare function, and maximizing over the choice of α , the regulator faces the following problem:

$$\max_{\alpha} W(q_1, q_2, \alpha) = B^T(q_1(\alpha)) + B^L(q_2(\alpha)) + \pi^*(q_1(\alpha), q_2(\alpha), K(\alpha))$$

where $\pi^*(q_1(\alpha), q_2(\alpha), K(\alpha))$ is the firm's maximized profit function (i.e the firm has already chosen outputs and the joint input). For this problem the first order conditions are:

$$W_{\alpha} = \frac{\delta B^T}{\delta q_1} \frac{dq_1}{d\alpha} + \frac{\delta B^L}{\delta q_2} \frac{dq_2}{d\alpha} + \frac{\delta \pi^*}{\delta q_1} \frac{dq_1}{d\alpha} + \frac{\delta \pi^*}{\delta q_2} \frac{dq_2}{d\alpha} + \frac{\delta \pi^*}{\delta K} \frac{dK}{d\alpha} = 0$$

Rearranging once again this yields:

$$W_{\alpha} = \left[\frac{\delta B^T}{\delta q_1} + \frac{\delta \pi^*}{\delta q_1} \right] \frac{dq_1}{d\alpha} + \left[\frac{\delta B^L}{\delta q_2} + \frac{\delta \pi^*}{\delta q_2} \right] \frac{dq_2}{d\alpha} + \frac{\delta \pi^*}{\delta K} \frac{dK}{d\alpha} = 0$$

(5.6.4)

Equation (5.6.4) shows that if a regulator knew that there would be additional effects of changing α on the firm's choice variables, then the regulator would choose α so that equation (5.6.4) is satisfied. Secondly, if regulators did not know about the effect of α on output and capital choice and chose α to maximize the "perceived" welfare using the maximization in equation (5.6.3), then the only way the outcome would be truly welfare maximizing (as defined by equation 5.6.4) would be if the following assumptions were true:

- i) $dK/d\alpha = 0$, and
- ii) $(\delta\pi/\delta q_1)(dq_1/d\alpha) = -(\delta\pi/\delta q_2)(dq_2/d\alpha)$

In this case regulators assume that capital is insensitive to changes in α and the effect of changing α on profits nets out. These are not implausible assumptions concerning what a regulator would know given the state of regulatory knowledge during the 1940's to the 1970's. Equation (5.6.4) yields a much different solution to the welfare agent's problem than the simple maximization of consumer surplus.

As mentioned above, the reasons why regulators could allocate costs from one jurisdiction to another during the past 40 years was due to the lack of substitutes for certain services and the availability of increasing returns to scale upon which the regulators could use to transfer rents between jurisdictions.

In addition, the structural benefits available to regulators when allocating costs have disappeared with the introduction of toll substitutes and the present existence of constant returns to scale in the toll market. Since 1969 when MCI entered the long distance market the number of substitutes from long distance service provided by AT&T has increased. There is also some evidence that (at least on the over all level) AT&T Long-Lines were operating at a point of constant returns to scale (Meyer: 1980). So, the factors which allowed regulators to shift costs from one jurisdiction to another based solely upon the firm's derived demand considerations are no longer present. Continuation of this type of cost allocation regulation poses additional dangers for regulators and the multiproduct firm as it could tend to undermine the competitiveness of the toll division of the multiproduct firm. This is the state of affairs from which the access charge arrangements must be developed.

5.7 Conclusion

From the first model in section 5.2 we found that a firm could select its own cost allocation (α) to maximize its profits given an average cost constraint. Further, if the regulators set the allowed rate of return equal to the cost of capital ($s = r$) in both jurisdictions and allowed the firm to choose output in both markets, capital

stock, and the cost allocator, the firm would pick a Ramsey price vector.

It was also shown that the firm's choice of the allocator could be influenced by the relative rates of return in each jurisdiction. If a firm was given an exogenous cost allocation mechanism and allowed to maximize profits given the choice of output and capital stock, a solution would exist consistent with the rate of return constraints. Comparative statics were then derived to show how the long run and short run decisions were affected by exogenous changes in the cost allocation mechanism. It was shown that in the long run that a change in the cost allocator could have a positive or negative effect upon the relative output mix and the choice of capital. In the short run the results were more definite as outputs changed in the expected manner when costs were shifted between jurisdictions.

Using Rosses' approach we demonstrated that it was possible to determine the optimal allocator based on just Ramsey price information, and it would be possible to test just how far the actual allocation differed from Ramsey optimality. Finally, using Joskow's approach we examined the welfare aspects of administratively allocating costs. We showed that by making simplifying assumptions concerning the effect of changing the cost allocation on the choice of input and output mix that regulators were ignoring import-

ant firm responses. By including the firm's responses into the analysis we show the proper cost allocation decision.

CHAPTER VI
CONCLUSIONS AND AREAS FOR FURTHER RESEARCH

This dissertation examined the cost allocation process in a relatively unique way. First, the capital input which was used to provide services for a multiproduct firm was described as a joint input and analyzed as a joint input. As was shown, the optimal choice of a joint input is determined by the sum of the marginal valuations for each demander for the joint input. Secondly, because we were using this framework it was easier to see that the choice of the cost allocator had an effect upon the overall level of the joint input selected by the firm.

Then by reviewing the history of the cost allocation process in the telecommunications industry, two main reasons for altering the allocation mechanism were discussed and the effects of administratively altering the cost allocation mechanism were examined. The first reason to alter the cost allocator, to promote universal service, was to use the allocator as a way of transferring income between markets to promote the subscription of local service for everyone. The second reason why the government desired to alter the allocation mechanism was to rectify imbalances between intrastate and interstate rates. By changing the cost allocation between markets the states would not have to grant increases in local rates.

Both of these reasons for changing the cost allocation mechanism had the underlying goal to transfer income between the two regulatory jurisdictions. Chapter IV and V showed that this transfer was not without distortions as the transfer effected the choice of the level of the joint input and the relative output mix. More specifically, Chapter IV showed, using a two product firm with fixed proportions production where one market had a technology experiencing increasing returns to scale, that there were costs to allocating more of the joint input to the market with economies of scale. Given that the firm is required to price at average cost, if the demand elasticity for the market with economies of scale was relatively elastic, then it was possible to show that a large reduction in output could occur. As price increased, due to the increase in costs resulting from the new cost allocation, demand would fall and the division would be operating at a higher point on the average cost curve than a division with constant returns to scale technology.

Chapter V dealt with a two product firm with a neoclassical production function where the production function of both markets uses a joint input. The object was to examine the firm's behavior under different regulatory assumptions. From this analysis it was determined that if a multiproduct firm was required to price at average cost in all markets and it was required to allocate

a share of the cost of the joint input to all divisions employing the input, then the firm would pick an allocator yielding Ramsey prices. If, however, the government was the agent determining the cost allocation mechanism and it altered the allocation for various policy reasons, then the change in the allocator could alter the amount of the jointly used input chosen (positively or negatively) and change the relative output mix of the firm. In addition, if it is supposed that the government's policy goal was to change the allocator to benefit a particular class of customers (say local customers) by transferring costs from local to toll, then only in the short-run would the government's policy be unambiguously effective. But if the change in the cost allocation effects the choice of the capital input, then the long run result of changing the cost allocation could have the exact opposite effect from what was intended. This result lead to the observation that regulators altered the cost allocation between interstate and intrastate markets based upon an incomplete understanding of the firm's technology or a lack of concern over the change in the firm's input choice or relative output mix.

A simple example of regulatory behavior using Joskow's passive regulation theory showed that different policy results occur given incomplete knowledge (e.g. regulators assume that a change in the cost allocation has no effect

on the level of profits) or a lack of concern for the firm's profit level as long as it was no higher than the regulatorily determined allowed rate of return. Also discussed was Rosses' (1984) assertion that it was possible to infer information concerning regulator's implied social welfare weights from the pricing decisions they made. Ross "extracted" implied social welfare weights from just the Ramsey solution. This analysis is extremely dependent upon the information known to the regulators. Chapter V pointed out that using a derivation identical to Rosses' it was possible to make alternative conclusions. Instead of extracting implied welfare weights it is possible to determine just how far away the cost allocation was from the allocation that yielded Ramsey prices. Rosses' interpretation implicitly assumes that regulators know what Ramsey prices or Weighted Ramsey prices are, and that regulators attempt to set prices based upon the Ramsey rule. Ross also implicitly assumes that assumes that the information concerning the elasticities and technology of the firm are known to the regulators and are correct.

Throughout this dissertation it was assumed that the goods produced by the multiproduct firm were independent in consumption. This was done for the sake of simplicity and because it made intuitive sense in the simple local-toll story that was being presented. It is important to note that when the goods are not independent there is a

more complicated Ramsey result that takes into account the cross elasticities of demand between goods. For example, if the two goods in question are substitutes one would expect that the Ramsey prices generated would be lower for each good than for the case where the goods are independent. It is also possible that the two goods could be complementary in consumption. For example, to access many national data bases a telephone subscriber must dial a local call. The call is then routed by a toll route to the data base. The subscriber only pays a local rate for this call, but actually uses the local network and the toll network (presumably the customer pays the cost of his use of the toll network through the data base service fee). As more and more services can be accessed through the telephone network and more and more uses are developed for the network, regulation through cost allocation changes will become very complicated. The effect upon the cost allocator under this more realistic scenario is something that should be examined.

Another area for further study concerns the effect of information on the regulatory decisions. Given the state of regulatory knowledge during the period from 1940 to 1970, it is quite plausible regulators assumed that profits were insensitive to changes in the cost allocation mechanism. In Chapter V it was shown that there exists an a relationship between the social welfare weight a regula-

tor may place on a particular good and the optimal cost allocation mechanism. A low cost allocation, for example, could thus imply a relatively higher social welfare weight. But, given the fact that the state of regulatory expertise was not very sophisticated until very recently, any derived weights would necessarily be the weights the regulators desired to place on these services. The "weights" derived, which would be the result of misinformation or the lack of information, could be used to see how far the regulator's policy deviated from a Ramsey solution. The literature concerning regulatory policies with imperfect information is relatively young (see Sappington: 1983) and a good empirical analysis would be to examine the differences between the regulator's perceived welfare maximizing policy and the true social welfare maximizing policies.

A third area of research into the telecommunications industry's cost allocation process would concern examining the interrelationship between the revenue distribution process, which is known as settlements, and the separations process. As the settlements process was originally set up, the amount of revenue received by the local operating telephone company to cover the costs of interconnection with the local network was a function of the nation-wide average rate of return on capital allocated by the separations formula to toll service multiplied by the amount of capital a firm and allocated to toll production. This had

the possible effect of penalizing efficient firms (those with below average costs) and rewarding inefficient firms (those with above average costs). Secondly the process resulted in the firm having a different objective function. With the toll settlements procedure the firm no longer maximizing profits, but maximized profits plus the return from a revenue pool. The revenue pool was funded by collecting revenues from all the long distance calls. After collection each local operating company then received its "share" which was determined by how much capital the local company had allocated, based on separations formula, to toll service. The interrelationship between separations and settlements and their joint effects on output mix and the capital input choice have yet to be explored.

Thirdly, the effect of administratively determined cost allocations on entry and sustainability needs to be analyzed. The traditional literature of entry and sustainability does not really address the issue of cost allocations. The new literature in this area centers around the theory of clubs as the allocation of the cost of a joint input can be thought of as a price for sharing the cost of a facility (see e.g. Scotchmer: 1985).

Finally, something that was briefly mentioned in Chapter III and by Panzar (1980) is the problem of allocating costs in a vertically integrated firm with economies of scope and where the firm also experiences competition in

one of its markets. This was the conjecture concerning the pre-divestiture telecommunications industry. The questions that arise in this area have not gone away after the consent decree because there are seven Little AT&T's in the form of the regional Bell operating companies that still produce some toll services. Presently some of the states are even thinking of allowing competition within the Bell operating company service area. The necessary research in this area must address the question: When should a vertically integrated firm, owning an input competitors desire to employ to compete against the vertically integrated firm, be allowed to exclude these competitors. The example in Chapter III pointed out that the multiproduct firm has little incentive to provide access to competitors; especially if the competition is allowed to purchase the right to use the joint input and share in the costs savings (due to economies of scope) and undercut the vertically integrated multiproduct firm. An interesting analysis would be to examine the welfare effects to see under what conditions the firm should be allowed to prohibit the competition from using its facilities and under what conditions the firm should be forced to allow access.

This dissertation has shown that there are some previously unrecognized effects of administratively altering the allocation of the cost of a joint input among the divisions of a regulated firm. Given the fact that

the access charge to be paid by toll interexchange carriers is for the support of the local network, regulators still must recognize that changes in this allocation to interexchange users will still have the same possible effects on the choice of the joint input and the relative output mix. Today, it seems that regulators are moving away from the fully distributed costing methodologies and are moving towards prices associated with long-run marginal costs, but when the time comes to determine the cost allocation process in the form of access charges, regulators should realize that politically determined changes in the mechanism could have more than the supposed redistributive effect.

BIBLIOGRAPHY

- Areeda, P., and D. F. Turner, "Predatory Pricing and Related Practices under Section 2 of the Sherman Act," Harvard Law Review 88 (1975), 697-733.
- Artle, R., and C. Avernous, "The Telephone System as a Public Good: Static and Dynamic Aspects," Bell Journal of Economics 4 (1973), 89-100.
- Averch, H., and L. Johnson, "Behavior of the Firm Under Regulatory Constraint," American Economic Review 52 (1962), 1052-1069.
- Auman, R. J., and L. S. Shapley, Values of Non Atomic Games, Princeton: Princeton University Press, 1974.
- Bailey, E., and A. F. Friedlander, "Market Structures and Multiproduct Industries," Journal of Economic Literature (1982), 1024-1148.
- Baumol, W. J., "On the Proper Cost Tests for Multiproduct Natural Monopoly," American Economic Review 67 (1977), 809-822.
- , and D. F. Bradford, "Optimal Departures from Marginal Cost Pricing," American Economic Review 60 (1970), 265-283.
- , J. Panzar, and R. D. Willig, Contestable Markets and the Theory of Industrial Structure, San Diego: Harcourt, Brace, Johanovich 1982.
- Berglas, E., and D. Pines, "Clubs, Local Public Goods, and Transportation Models: A Synthesis," Working Paper 32/78, Foorder Institute for Economic Research, Ramat-Aviv, Isreal, October 1978.
- Billera, L. J., D. C. Heath and J. Raanan, "Internal Telephone Billing Rates," Operations Research 26 (1978).
- Braeutigam, R. R., "An Analysis of Full Distributed Cost Pricing in Regulated Industries," Bell Journal of Economics, 11 (1980), 182-196.

- Cohen, S. I., and M. Loeb, "Public Goods, Common Inputs, and the Efficiency of Full Cost Allocations," Accounting Review 57 (1982), 336-347.
- Dougherty, A. N., "Emprical Estimates of Demand and Cost Elasticities of Telephone Service," Changing Patterns in Regulation, Markets, and Technology, East Lansing: MSU Institute of Public Utilities 1984, 115-141.
- Ellickson, B., "Public Goods and Joint Supply," Journal of Public Economics 9 (1978), 373-382.
- Faulhaber, G. R., "Cross Subsidization: Pricing in Public Enterprises," American Economic Review 65 (1975), 966-977.
- Fumas, V. S., and A. B. Whinston, "Subsidy Free Welfare Games," Southern Journal of Economics 49 (1982), 389-405.
- Gabel, R., History of Separations, East Lansing: MSU Public Utilities Institute 1967.
- Hamlen, S., W. Hamlen, and J. Tschirhart, "The Use of Core Theory in Evaluating Joint Cost Allocation Schemes," Accounting Review 52 (1977), 626-627.
- Hannon, J. T., "The Impact of Competition on Cost Allocation in a Multi-market Monopoly: Telecommunications," Ph.D. dissertation, University of Illinois at Champaign Urbanna, 1978.
- Heany, J. P., and R. E. Dickenson, "Methods for Apportioning the Costs of a Water Resources Project," Water Resources Research 18 (1982), 476-482.
- Kahn, A. E., Economics of Regulation, New York: Wiley 1971.
- _____, "The Road to More Intelligent Telephone Pricing," Yale Journal on Regulation 2 (1984), 139-158.
- Lavey, W. G., and D. W. Carleton, "Economic Goals and Remedies of the AT&T Modified Final Judgement," Georgetown Law Review 71 (1983), 1497-1518.
- Leggette, J., "Separations and Telecommunications Pricing Issues," forthcoming in Proceedings of the Georgia Public Utilitiy Conference in press.

- Lindahl, E., "Just Taxation--A Positive Solution," in Classics in Theory of Public Finance, R. A. Musgrave, editor, London: Macmillan 1958.
- Littlechild, S. C., and G. F. Thompson, "Aircraft Landing Fees: A Game Theory Approach," Bell Journal of Economics 8 (1977), 186-203.
- Loeman, E. and A. Whinston, "A New Theory of Pricing and Decision Making for Public Investment," Bell Journal of Economics 2 (1971), 606-627.
- Marshall, A., Principles of Economics London: MacMillan 1927.
- MacAvoy, P. W., and K. Robinson, "Winning by Losing: The AT&T Settlement and Its Impact on Telecommunications," Yale Journal on Regulation 1 (1983), 1-42.
- Miriman, L. J., D. Savvet, and Y. Tauman, "An Axiomatic Approach to the Allocation of Fixed Costs through Prices," Bell Journal of Economics 13 (1983), 139 - 151.
- Moriarty, S., "Another Approach to Allocating Joint Costs," Accounting Review 50 (1975), 791-795.
- Needy, C., Regulation Induced Disortions, Lexington: Lexington Books 1975.
- Oettinger, A., The Federal Side of Telecommunications Cost Allocations, Harvard University Program on Information Resources Policy Paper (1980).
- Ordover J., and R. Willig, "Anticompetative Practices by Dominant Firms," Yale Law Journal 91 (1982).
- Panzar, J.C., "Sustainability, Efficiency, and Vertical Integration," in Regulated Industry and Public Enterprise ed. by B. M. Mitchell, and P. R. Kleindorfer, Lexington: Lexington Books 1980, 171-203.
- _____, and R. D. Willig, "Free Entry and the Sustainability of Natural Monopoly," Bell Journal of Economics 8 (1977), 1-23.
- Pelcovits, M. D., N. W. Cornell, and S. R. Brenner, "A Slip twist Cup and Lip," Changing Patterns in Regulation, Markets, and Technology, East Lansing: MSU Institute of Public Utilities 1984, 343-359.

- Pigou, A. C., "Railway Rates and Joint Costs," Quarterly Journal of Economics 27 (1913), 687-692.
- Porter, O.W. "History of Settlements," presentation made June 1979 (mimeographed).
- Ramsey, F., "A Contribution to the Theory of Taxation," Econometric Journal 37 (1927), 47-61.
- Ross, Thomas, "Uncovering Regulators' Social Welfare Weights," working paper of Center for the Study of the Economy and the State, University of Chicago (1982).
- , "Uncovering Regulators' Social Welfare Weights," Rand Journal of Economics 15 (1984), 152-155.
- Samuelson, P. A., "The Pure Theory of Public Expenditures," Review of Economics and Statistics, 36 (1954), 387-389.
- , "Pure Theory of Public Expenditures and Taxation," Public Economics ed. by J. Margolis and H. Guitton, New York: St. Martin's 1969.
- , Foundations of Economic Analysis, Cambridge: Harvard University Press 1983.
- Sappington, D., "Optimal Regulation of a Multiproduct Monopoly with Unknown Technological Possibilities," Bell Journal of Economics 13 (1982), 354-368.
- Sandler, T., and J. Tschirhart, "The Economic Theory of Clubs: An Evaluative Survey," Journal of Economic Literature 28 (1980), 1481-1520.
- Schotter, A., and G. Schwodiauer, "Economics and the Theory of Games: A Survey," Journal of Economic Literature 28 (1980), 479-527.
- Scotchmer, S., "Two Tier Pricing of Shared Facilities in a Free Entry Equilibrium," Rand Journal of Economics 16 (1985), 456-472.
- Shapley, L., "The Value of an N-Person Game," in H. W. Kuhn and A. W. Tucker (eds.), Contributions to the Theory of Games Princeton: Princeton University Press, 1953.
- Sharkey, W. W., "Existence of Sustainable Prices for Natural Monopoly Outputs," Bell Journal of Economics 12 (1981), 144-154.

- _____, "Suggestions for a Game Theoretic Approach to Public Utility Pricing and Cost Allocation," Bell Journal of Economics 13 (1982a), 57-65.
- _____, The Theory of Natural Monopoly, New York: Cambridge 1982b.
- Shavell, S., "Risk Sharing and Incentives in the Principal-Agent Problem," Bell Journal of Economics 10 (1977), 175-188.
- Sichter, J. W., Separations in the Telephone Industry, Cambridge: Program on Information Resources Policy, Harvard University 1977.
- Sickler, B. J., "A Theory of Telephone Rates," Journal of Land and Public Utility Economics 6 (1928), 175-188.
- Sparling, L. T., "Regulatory Distortions in Transportation and Telecommunications," Ph.D. dissertation, California Institute of Technology 1980.
- Stein, J. L., and G. H. Borts, "Behavior of the Firm under Regulatory Constraint," American Economic Review 42 (1972), 964-970.
- Taussig, F. W. "A Contribution to the Theory of Railway Rates," Quarterly Journal of Economics 5 (1891), 535 - 538, and 687 - 694.
- Temin, P. and G. Peters, "Cross-Subsidization in the Telephone Network," Williamette Law Review in press.
- _____, "Is History Stranger than Theory?: The origins of Telephone Separations," American Economic Review 75 (1985), 324-327.
- Wellisz, S. H., "Regulation of Natural Gas Pipeline Companies: An Economic Analysis," Journal of Political Economy (1963), 30-43.

Legal Bibliography

- Above 890 MHz, 27 FCC2d 357 (1962).
- Communication Act of 1934, 47 U.S.C. § 101 et seq. (1982).
- Second Computer Inquiry (Computer II), 88 FCC2d 512 (1981)

Exchange Network Facilities for Interstate Access (ENFIA),
71 FCC2d 440 (1979).

FPC v. Hope Natural Gas 320 U.S. 591 (1944).

MCI v. FCC, (Execunet I) 561 F.2d 365 (D.C. Cir.,1977)

MCI v. FCC, (Execunet II) 580 F.2d 590 (D.C. Cir. 1977).

Microwave Communications, Inc., 18 FCC 2d 953 (1969).

Minnesota Railroad and Warehouse Commission v. Shephard,
230 U.S. 352 (1913).

NARUC v. FCC, no. 83-1225 (D.C. Cir. 1984).

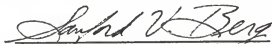
Smith v. Illinois Bell, 282 U.S. 133 (1934).

U.S. v. AT&T Co., 524 F. Supp. 1336 (D.D.C. 1981).

BIOGRAPHICAL SKETCH

Martin F. Grace, born in El Paso, TX, in 1958 lived throughout the U.S. as a youth. He studied political science at the University of New Hampshire, graduating in 1980. After working briefly at the Federal Communications Commission, Mr. Grace commenced graduate study at the University of Florida and received a M.A. in economics in 1981. After finishing the M.A., Mr. Grace enrolled at the University of Florida College of Law. A doctoral degree is expected in December 1985 and a law degree is expected in December 1986.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Sanford V. Berg, Chairman
Associate Professor of Economics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



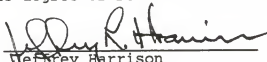
Roger D. Blair
Professor of Economics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Virginia Wilcox-Gök
Assistant Professor of Economics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.



Jeffrey Harrison
Professor of Law

This dissertation was submitted to the Graduate Faculty of the Department of Economics in the College of Business Administration and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 1986

Dean, Graduate School